Buzz vs. Sales: Big Social Data Analytics of Style Icon Campaigns and Fashion Designer Collaborations on H&M's Facebook Page

Desiree Christina Komtesse af Rosenborg [1], Ida Buhl-Andersen[1], Line Bygvrå Nilsson[1], Mark Philip Rebild[1], Raghava Rao Mukkamala[1] Abid Hussain [1] and Ravi Vatrapu[1,2]
[1]Copenhagen Business School, Denmark and [2] Westerdals Oslo School of ACT, Norway
{rrm.itm, ah.itm, rv.itm}@cbs.dk

Abstract

This paper examines the relationship between social media engagement and financial performance of the global fast fashion company, H&M. We analyze big social data from Facebook on the seven H&M style collections that occurred during 2012 and 2013 to investigate if style icon campaigns have a larger effect on quarterly sales than designer collaborations. We find that style icons such as David Beckham generate more social buzz than designer collaborations. Social Set Analysis of the Facebook data shows that the overlap between the users H&M reach with their different style collections is fairly small. The deviations between forecasted quarterly sales and actual quarterly sales are analyzed. Our results show that that style icon campaigns have a larger impact on sales than designer collaborations and reveal that the quarters with the largest deviations coincide with the quarter in which H&M ran a style icon campaign. We discuss the implications of our findings and outline directions for future research.

1. Introduction

Much human activity now takes place online, which generates massive amounts of data with substantial commercial value [1]. Especially the use of social media has exploded in the last couple of years, resulting in massive amounts of social data. Big data analytics has, among other things, the power to transform the way firms conducts business [2]. This paper explores social data from H&M's Facebook page in relation to specific marketing campaigns initiated by H&M as well as in relation to the financial performance of H&M. First, we investigate the social activity around selected style and designer collections. Second, we explore the relation between the social activity and quarterly financial sales. Using modeling forecasted sales and visual analytics, we seek to explain deviations in actual sales versus forecasted sales in relation to the social media activity around these campaigns.

H&M is a Swedish, high street fashion retail company engaging in the sale of clothing, accessories, footwear, cosmetics, and home textiles worldwide. The company's brands include COS, &Other Stories, Weekday, Cheap Monday, Monki, and H&M Home as well as the H&M brand[1] (H&M 2). One of the reasons why H&M was found relevant for the analysis was due to the fact that it is a company that has substantial social interaction on their Facebook page. Second, H&M has over the last couple of years continuously initiated campaigns within certain areas, hence this made it possible for us to analyze the development and patterns in the social activity around these campaigns and investigate whether there was an impact between the social activity around the campaigns and the financial performance. The campaigns by H&M is based on different style collections, which can be divided into three overall categories; style icons, designer collaborations and *good causes*. Throughout the paper the term *style collection* will be used as a common term for style icon campaigns and designer collaborations.

1.1 Research Questions

Our primary objective is to investigate H&M's style collections and analyze their performance as well as their relationship with sales. Our research question is:

> *How and to what degree does H&M's style collections affect social media performance and quarterly sales?*

In order to answer the above, we formulated one main proposition and five sub-propositions which will be empirically analyzed. Based on the fact (table 1) that style icons are much more popular on social media than fashion designers, our main proposition is that: *Style icon campaigns perform better on social media than designer collaborations*. The style icon campaigns consist of Beyoncé, David Beckham and Lana Del Rey and the designer collaborations consist of Versace, Maison Martin Margiela (MMM), Marni, and Isabel Marant. Our

[1] H&M Campaigns and designer collaborations

Style Icon Campaigns			Designer Collaborations			
Beyoncé	David Beckham	Lana Del Rey	Versace	MMM	Marni	Isabel Marant
63.791	53.760	11.904	4.094	0.352	0.174	0.113

Table 1: Facebook Page Likes (in Millions)

propositions aims at investigating the social data combined with business data in detail to identify social buzz around each style collections.

- **Sub-proposition 1**: On average, style icon campaigns result in more social buzz than designer collaborations
- **Sub-proposition 2**: Style icon campaigns will have more likes per post than designer collaborations, since a style icon can reach a broader audience than a fashion designer
- **Sub-proposition 3**: H&M is more aggressive in their Facebook promotion of designer collaborations than style icon campaigns
- **Sub-proposition 4**: Style icons target a more differentiated audience than designer collaborations
- **Sub-proposition 5**: In a sales forecast, the quarters with style icon campaigns will have the largest deviations between actual and forecasted sales

2. Related Work

There is elaborate body of work done on predictive analytics with Big Social Data and the most relevant related work for our paper is highlighted below.

2.1 Data Science Perspective

There has been substantial research work [3–8] in the direction of predicting the stock prices of the companies based on the analysis of content from the online media such as news items, web blogs, twitter feeds. For example, Gavrilov et al., [6] applied data mining techniques on the stock information from various companies by clustering them according to their Standard and Poor (S&P) 500 index, whereas the content from the weblogs is used by Kharratzadeh & Coates [7] to identify the underlying relationships between the companies to make predictions about the evolution of stock prices. Asur & Huberman [9] showed that social media feeds can be used as effective indicators of the real-world performance. In their work, they used analysis of hourly rate of tweets about movies, their re-tweets and sentiment polarity to accurately forecast the box-office movies revenue. Lassen and colleagues predicted iPhone sales from iPhone tweets [10]. In terms of macro-societal relationships, a research study investigated whether the public mood as measured from large-scale collection of Twitter tweets can be correlated or even predictive of Dow Jones Industrial Average (DJIA) values has been explored by Bollen and Mao [4].

2.2 Information Systems Perspective

Seebach et al. [11] suggested that companies include data on customer's online search into their IT systems in order to increase their sensing abilities and create a more agile business. Moreover, Reijden & Koppius [12] studied how online buzz predicts actual sales across different phases of a product lifecycle. Geve and colleagues [13] used Google's index of internet discussion forums and Google's search trends to predict sales, while Wu and Brynjolfsson [14] used internet searches to predict housing prices. Zhang and Lau [15] developed a business network-based model to analyze and predict business performances (using the proxies of stock prizes). Nann, Krauss, and Schoder [16] analysed multiple online public data platforms such as Twitter and Yahoo! Finance in order to predict the stock market, while Oh and Sheng [17] analysed the predictive power of micro blog sentiments on stock price directional movements. The authors Hoffman and Fodor [18] advocated that proper social media strategies developed in the context of 4c's (connections, creation, consumption and control) can lead to higher return on investment on their marketing investments and also leads to better active customer engagement. Another research [19] studied the role of social media in sales and marketing and they argued that the future of social media in sales and marketing depends both on the customer and the company, where they both evolve with the time.

When compared to above related work, our approach differs in the sense that we study the effect of style collections on the quarterly sales by using event study and social set analysis methodologies to compute relationship between the social data (e.g. Facebook data) and financial performance (e.g. quarterly revenues).

3. Conceptual framework

Social buzz refers to the engagement created by consumers talking about a company's products on social platforms. Following the idea of buzz marketing, by getting consumers to talk about their products and services, companies can succeed in creating a social buzz. Thereby they are more likely to increase awareness of their products because of growth in online traffic leading to increased sales and profits[2]. In our analysis, social buzz will be operationalised and measured as the social media engagement in terms of the quantity of posts, comments and likes on the official Facebook page of H&M.

Brand ambassadors are the focal point of the campaigns analyzed throughout this paper. Brand ambassadors are typically used in marketing to increase atten-

[2] Buzz Marketing

tion, interest and desire around a brand or a brand collection. Since celebrities and fashion designers already have an established image and identity in public, using them as brand ambassadors helps express a clear identity of the brand, making it easier for the consumer to identify with the brand [20]. Forecasting methods are typically grouped into two groups, static and adaptive models. Static methods assume that estimates of level, trend, and seasonality do not vary as new values for sales are observed whereas adaptive forecasting updates the estimates after each new observation [21]. As the historical data for our forecast is limited to only 3 years we wish to apply an adaptive forecasting method to include as much information as possible for each consecutive quarter.

4. Research Methodolgy

The datasets used for the analysis are: designer and style icon events, social data from the H&M's Facebook page and financial data of H&M.

4.1 Dataset Description

4.1.1. Style Icons and Designer Collections Data Due to the focus on the different campaigns by H&M it is relevant to examine the marketing strategy behind these initiatives further. H&M's designer collaborations have all been launched as limited editions at exclusive design shows with the presence of top models and high profile people[3]. Besides the designer collaborations, numerous pop and style icons have featured in the H&M campaigns. Both categories of campaigns are promoted on Facebook. Overall, these two categories of campaigns involve:

1. Style Icons: Campaigns where a celebrity (not related to fashion) promotes a collection.
2. Designer collaborations: Campaigns collaborated with a famous fashion designer, either within haute couture or exclusive fashion brands. These designers are different from normal H&M's target group, which makes the collaborations rather unique.

The interesting campaigns within these two categories during 2012 and 2013 are shown in table 2. Our interest is to explore performance of these campaigns and their impact on overall performance of H&M.

4.1.2. Social Data The raw dataset consisted of just above 13 million data points. Each row is equivalent to an action on H&M's Facebook page. The data are ordered into dimensions, which include both ordered and categorical data and contains information about actor ID, timestamp, event name (whether it is a post, comment, like), actor name, type of post and if relevant links and text value of post and comments. The social data available goes from 2009 until the third quarter of 2014, however, for the present analysis only social data from 2012 and 2013 have been relevant. The aggregated data of total posts, comments and likes both for H&M, non-H&M as well as combined, as shown in table 3.

Year	AQ[4]	RQ	Style Collection Type	Style Collection	Keywords
2012	Q2(11)	Q1	Designer	Versace Cruise	Versace
2012	Q2(11)	Q1	Style icon	David Beckham	David Beckham
2012	Q3(11)	Q1	Designer	Marni	Marni
2012	Q1	Q2	Aids	Fashion Against Aids	Aids
2012	Q3	Q3	Style icon	Lana Del Re	Lana Del Re
2012	Q2	Q4	Designer	Maison Martin Margiela	Maison Martin Margiela
2013	Q1	Q2	Style icon	Beyoncé	Beyoncé
2013	Q2	Q4	Designer	Isabel Marant	Isabel Marant

Table 2: Overview of style collections during 2012-13

	Versace	David Beckham	Marni	Lana Del Re	MMM	Beyoncé	Isabel Marant
Total							
Posts	1,321	232	926	193	371	292	982
Comments	2,473	4,548	1,175	2,449	1,603	4,447	3,128
Likes	37,120	493,904	25,713	128,707	68,576	192,203	191,722
Before the Event							
Posts	1,078	74	557	126	198	277	274
Comments	2,051	900	764	1,632	440	4,429	1,408
Likes	34,585	21,295	18,817	74,180	27,590	191,942	95,873
During the Event							
Posts	179	99	344	41	138	2	668
Comments	348	2,393	368	588	1,109	1	1,645
Likes	1,575	63,331	6,673	42,557	40,525	1	94,266
After the Event							
Posts	64	59	26	24	33	12	40
Comments	74	1,255	43	229	54	17	75
Likes	960	409,278	226	11,970	461	260	1,583

Table 3: Aggregated data of Style Icons and Designers

4.1.3. Business Data The financial data H&M quarterly was collected from the annual reports and limited to the periods of 2009 until the third quarter of 2014 like the social data.

4.2 Data Analytics Process

The data analysis is illustrated in figure 1. The social data set was collected using a tool SODATO [22, 23] We have partitioned the data using specific relevant keywords for campaign and ended up seven individual datasets. Then the data is separated into categories, by creating a categorical variable for each campaign, which is used in further analysis and visualizations. The keywords used for identifying relevant data points are shown in table 2.

[3] H&M's 5 Modes Of Collaboration

[4] AQ = Announced Quarter, RQ = Collection Release Quarter

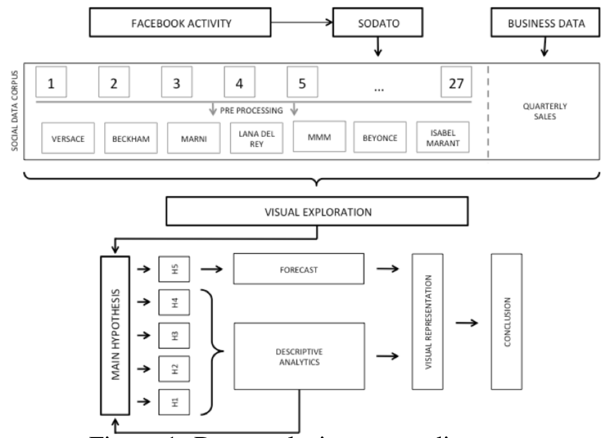


Figure 1: Data analysis process diagram

In order to calculate statistical distributions of each campaign, we classified data based actions done by H&M or non-H&M Facebok users as shown in table 4. Combining different data sources will create huge challenges as well as opportunities for unlocking great value [24]. By combining sales data and social data from Facebook, insights about social buzz for specific campaigns can be identified and their impact on sales can be assessed.

	Versace	**David Beckham**	**Marni**	**Lana Del Re**	**MMM**	**Beyoncé**	**Isabel Marant**
				H&M Activity			
Posts	20	15	11	13	0	4	24
Comments	193	25	56	24	5	40	87
Likes	40	11	17	16	0	11	19
				Non-H&M Activity			
Posts	1,301	217	915	180	371	288	958
Comments	2,280	4,523	1,119	2,425	1,598	4,407	3,041
Likes	37,080	493,893	25,696	128,691	68,576	192,192	191,703

Table 4: H&M & non-H&M activity for each campaign

4.3 Data Analytics: Predictive Modelling

Many adaptive forecasting methods are available: moving average, simple exponential smoothing, Holt and Winter's forecasting method [21]. However, a suitable method for for H&M sales data will be chosen based analysis. The assessment relies on the degree to which the trend (the rate of growth/decline in demand), seasonality (predictable fluctuations in demand) and level (the deseasonalized current demand) aspects are present in the sales time-series data.

Observing the development of sales clarifies how a trend of increasing sales is affecting it. Without exception the sales in the first quarter of a year is higher than in the same quarter of the previous year. Our choice of forecasting method must thus be capable of incorporating this into the forecast for us to be able to measure the effect of campaigns. At closer examination it is also clear how seasonality heavily influences the sales development. With the exception of the third and fourth quarter of 2013, the other quarters consistently experienced an increase in sales while the first and third quarter of a year consistently experienced a decrease in sales compared to the prior quarter. Therefore, our forecasting method should also be capable of incorporating seasonality into its forecasts. As neither the forecasting methods of moving average nor simple exponential smoothing is capable of incorporating trend or seasonality to their forecasts [21], these methods are excluded. In addition, Holt's method is unable to incorporate seasonality. We are thus forecasting the sales for H&M for years 2012 and 2013 using the Winters method which includes both trend and seasonality, thereby allowing us to measure the impact of the style icon and designer collaboration campaigns.

4.3.1. Reasoning for Smoothing Constants Characteristic of the Winters method is that none of the estimates remain constant over time. After each period of observed sales they are changed to include part of the information contained in the previous quarter through the use of smoothing estimates for alpha, beta, and gamma. Alpha is used to smooth the level, beta the trend, and gamma the season. In general, it is advisable to use smoothing constants around 0.2 or lower [21]. On occasion the constant can be increased for a short amount of time, but doing so makes the forecasts more sensitive to recent data. Thus, a better understanding of the underlying pattern in a data series should result in a lower smoothing constant than otherwise. Furthermore, the choice of smoothing constants can be based on whatever level minimizes the errors associated with the forecast. As we wish to include recent sales data for both level and trend at a comparable level, the values of alpha and beta are close to each other yet not equal. Based on the development from year 2009 to 2011 the trend seems more stable than the level, as the level fluctuated quite a bit in year 2011. Based on this we set the estimates for alpha to 0.15 and beta to 0.20. Based on the historical development of H&M sales we believe the seasonality to continue to be very constant through the years 2012 and 2013 and feel no need to smooth our estimate. Thus, we set gamma to zero and conduct our forecast with the same value for all quarters.

4.4 Forecasting Method

Notation: We use L_t for *level* at a time t, T_t for *trend* at time t, S_t denotes seasonal factor at time t and finally D denotes deseasonalized sales.
Using Winters method of forecasting, first we estimate the current level, L_T, the trend, T_T, and the seasonal factor for each quarter in order for us to calculate the systematic component. This calculation excludes the random component not explainable by historical data or underlying patterns related to trends or seasonality. Using [21], the

systematic component measures of the expected value of sales for H&M is calculated using a static method: *Systematic component = (level + trend) X seasonal factor*. Using static forecasting method, a forecast at time $t+l$ is based on the forecast at time t

$$F_{t+l} = [\, L + (t+l)T \,] \times S_{t+l} \tag{1}$$

where L = deseasonalized estimate at time period $t = 0$, T = estimate of trend i.e. increase or decrease per time period, S_t = estimate of seasonal factor for period t.

In order to estimate level L, trend T, the first step is *deseasonalizing* the actual sales data, which is the sales that would have been observed in the absence of seasonal fluctuations. Assuming that p is the periodicity, the number of periods after which seasonal cycle repeats, the deseasonalized sales $\overline{D_t}$ can be calculated using following equation from [21].

$$\overline{D_t} = \begin{cases} \left[D_{t-(p/2)} + D_{t+(p/2)} + \sum\limits_{i=t+1-(p/2)}^{t-1+(p/2)} 2D_i\right] \Big/ 2p & \text{if } p \text{ even;} \\ \sum\limits_{i=t-(p/2)}^{t+(p/2)} D_i/p & \text{otherwise.} \end{cases}$$

In order to ensure the equal weight has been given while calculating the deseasonalized sales, the average of sales for p consecutive periods is taken as shown in the above equation. The average sale from period $l+1$ to $l+p$ provides a deseasonalized sale value for the period $l+(p+1)/2$. In case if p is odd, the above equation provides deseasonalized sale for an existing period, alternatively, If p is even, then it provides deseasonalized sales at a point between Period $l+(p/2)$ and $l+1+(p/2)$. Therefore we take average of deseasonalized sales calculated for the periods: $l+1$ to $l+p$ and $l+2$ to $l+p+1$ to get an deseasonalized sales for period $l+1+(p/2)$. As we are dealing with quarterly sales data in the case of H&M, we have adopted $p = 4$. Then we calculate seasonal factor S_t for the period t as the ratio of actual sales D_t to deseasonalized sales $\overline{D_t}$ as: $S_t = \frac{D_t}{\overline{D_t}}$.

4.4.1. Smoothing We now turn our attention back to the smoothing of the estimates to find out how it affects our forecast. Considering periodicity of sales as p and initial estimates of level (L_0), trend (T_0), seasonal factors ($S_1, \ldots S_p$), and in period t estimates of level (L_t), trend (T_t), seasonal factors ($S_t, \ldots S_{t+p-1}$), the forecasts for future periods are given by:

$$F_{t+1} = (L_t + T_t)S_{t+1} \text{ and } F_{t+l} = (L_t + lT_t)S_{t+l}$$

On observing sales price for Period $t+1$, we can revise the estimates for level using the following equation:

$$L_{t+1} = \alpha(D_{t+1}/S_{t+1}) + (1-\alpha)(L_t + T_t) \tag{2}$$

By using an α value of 0.2 we are including 20% of the information of the most recent observation of sales while relying 80% on our previous estimate of the level and trend. It is thus a weighted average of new information and estimated values. To properly incorporate the most recent information we deseasonalized it by dividing recent sales by the seasonal factor, leaving only the systematic component remaining. This way the estimate for the level is continuously updating and affected by the most recent information. Not only does this keep our estimates relevant, it also broadens the otherwise rather slim base of data for our forecasts. The trend is smoothed via an estimate for beta as shown in following equation:

$$T_{t+1} = \beta(L_{t+1} - L_t) + (1-\beta)T_t \tag{3}$$

By using a β value of 0.15 we are calculating an estimate for our trend based 15% on the most recent development in the estimated level of sales and 85% on the trend level calculated in the previous period. Once again we are basing our most recent estimates partly on the most recent data available to us, thereby keeping the estimate as relevant as possible while also increasing the amount of data used for forecasting. Finally, the seasonality is smoothed via an estimate for gamma and is calculated by the following equation:

$$S_{t+p+1} = \gamma((D_{t+1}/L_{t+I}) + (1-\gamma)S_{t+1} \tag{4}$$

For illustrative purposes, a gamma estimate of 0.1 would calculate an estimate for the seasonal factor based 10% on the ratio of actual sales to estimated sales in the most recent period and 90% on the most recent estimated value for the seasonal factor.

4.4.2. Forecasting of H&M Sales First the deseasonalized sales are calculated followed by seasonal factors for each of the four quarters as described in section 4.4. We have also conducted the same forecast using the Holt's method and the calculations for Holt's method are identical to those of Winters. Based on comparison of *mean square error* (MSE) of the two methods, Winters method found to be more superior and hence we proceeded with Winters' forecast for this paper. Figure 2 shows actual sales (2009-2011) and forecast (2009-2013). Though not perfect in explaining the development of sales, our forecast captures the level, trend and seasonality to a satisfying degree. The forecast for 2012 and 2013 will serve as the benchmark for the success of campaigns by attempting to explain the variances from actual development of sales through means of big social data analytics.

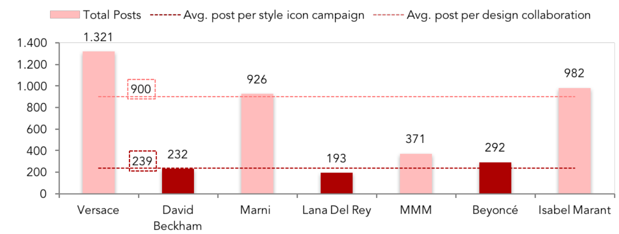
5. Findings and Results

In this work, we investigate if H&M's style icon campaigns perform better than those by designer collab-

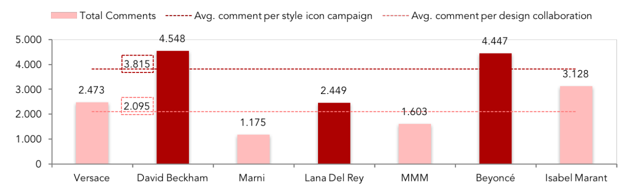
Figure 2: Actual sales verses forecasted sales

orations. We assume that style icon campaigns target a broader audience than designer collaborations through a wider fanbase and therefore create more social buzz leading to more sales. The correlation between social buzz and sales is believed to be proportional and positive. Our first set of propositions concern social buzz with the analytical focus on the nature and magnitude of the campaigns, where as final proposition concerns customer segments and financial performance. In this section, we will evaluate each of the proposition and conclude whether it is *supported* or *not supported* based on empirical findings. Our main proposition is *H&M's style icon campaigns perform better than designer collaborations*.

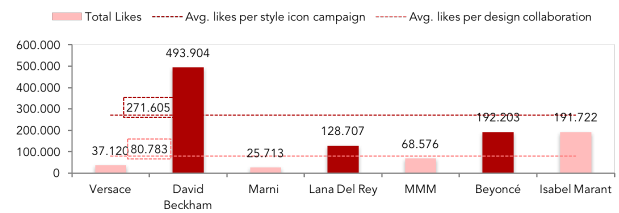
P_1-On average, style icon campaigns create more social buzz than designer collaborations: With the degree of social buzz defined as the amount of posts, comments and likes, P_1 helps evaluate our main proposition. This is tested by the total number of posts, comments and likes that each campaign got during its lifetime and their respective averages for all style icons campaigns and designer collaborations as shown in figure 3. The average number of posts for an H&M style icon campaign is 239. Contrary to the proposition, the average H&M designer collaboration get 276% more posts (900 posts) than a style icon campaign. This is a somewhat surprising, as we observed a significant and stable difference across campaigns and collaborations. It might be, because a style icon is not an *artistic creator* in the same way as a fashion designer is, hence it can be argued that there are stronger opinions about fashion designers than about style icons. Furthermore, an H&M style icon campaign gets on average 3,815 comments whereas a designer collaboration gets on average 2,095. This means the style icon campaigns gets about 82% more comments than collaborations, which is in line with the proposition. Similarly, in line with sub-proposition 1, style icon campaigns got 236% more likes than designer collaborations. Nevertheless it should be noted that it is primarily the David Beckham campaign that drives up the average number of likes for style icon campaigns, and without this extreme the difference between average



(a) Total and average number of posts for each style collection



(b) Total and average number of comments for each style collection



(c) Total and average number of likes for each style collection

Figure 3: Aggregated posts, comments and likes

likes per style icon campaign and designer collaborations would be a lot less. It is also interesting to note that over time designer collaborations have managed to generate more and more likes. Comparing the averages of respectively posts, comments and likes we find:

- Posts:designer collaborations generate most posts $\implies$ P_1 *not supported*
- Comments: style icon campaigns get most comments $\implies$ P_1 *supported*
- Likes: style icon campaigns get more likes $\implies$ P_1 *supported*

Thus P_1 is not completely supported, but we conclude that style icons create more social buzz, in terms of comments and likes than designer collaborations. The conclusion is relatively solid, since all activity around the specific campaigns of interest has been analyzed.

P_2-Style icon campaigns will on average have more likes per post than designer collaborations, since a style icon can reach a broader audience than a fashion designer, which mostly reach fashion savvy users: Because more Facebook users can be reached by having a broader audience (table 1), style icons should create more social buzz and consequently P_2 help us test our main proposition. We measure the broadness by how many likes H&M's icons get on average per post compared to H&M's designer collaboration. The likes to post ratio is calculated for each style collection by dividing the total number of likes for each style collection by the total number of posts that the style collection generated. Figure 4 illustrates the seven ratios and the average for style icons and designer collaborations and one could notice a significant difference between them. This difference is noteworthy even if the extreme values

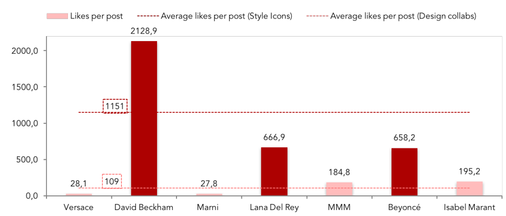


Figure 4: Likes per post for each style collection

of David Beckham was removed. It is apparent that style icons generate more social buzz than designers - and do it more efficiently, with fewer posts required to reach people. Thus P_2 is supported. Furthermore, it was noticed that when the style icon posted something about the campaign on their own Facebook page, they got quite a large amount of likes from just one post, due to a huge fan base on Facebook.

Finally, it was also noted that photo posts usually get more likes. The difference between likes per post for designer collaborations and style icons might therefore also be explained by the fact that style icons is functioning as a brand ambassador, provide a familiar and easily recognizable face in a photo post. In contrast, a photo of a dress designed by a famous designer might not be as easy to recognize, and hence needed description of posts.

P_3-H&M are more aggressive in their Facebook promotion of designer collaborations than style icon campaigns: As it is assumed (main proposition), that

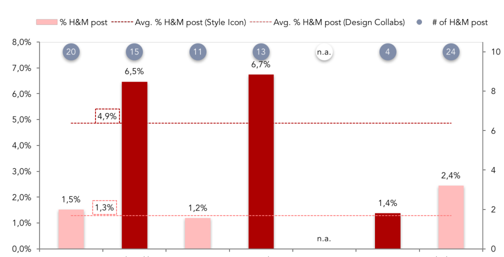


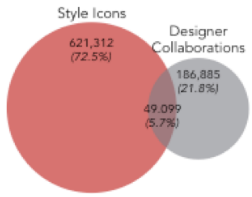
Figure 5: Post ratios for each style collection

style icon campaigns naturally create more social buzz, it is also anticipated that H&M will generate more posts for designer collaborations, to make up for the lower social buzz. P_3 will be tested by calculating how big a percentage of posts created by H&M make up the total number of posts for each style icons campaign and designer collaboration, the results are shown in figure 5.

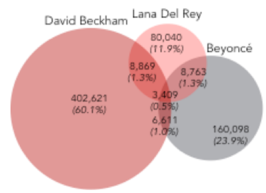
The results are somewhat surprising, as it looks like H&M make more posts about their style icons than their designer collaborations. Thus P_3 is not supported. However in absolute terms H&M actually generate more posts for designer collaborations on average (18 posts) than style icon campaigns (11 posts). Nonetheless the result is still unexpected because it would be expected that a style icon campaign would drive itself organically, meaning, in relative terms, H&M's post should make less number of total posts. One would believe that less famous designers would "need" more promotion than the more "universal" icons, but it is not the case.

P_4-Style icons target a more differentiated audience than designer collaborations: To understand relationship between audiences of the different style collections, we have employed the method of Social Set Analysis [25] and created three different Venn diagrams. Figure 6(a) shows the overlap between users who have acted (posted, commented or liked) on both style icon campaigns and designer collaborations. Figure 6(b) illustrates the overlap between users who have acted on the different style icon campaigns and the third diagram in figure 6(c) grabs the overlap in the users' interest in the different design collaboration again measured by act (post, comment, like). These Venn diagrams will therefore indirectly help us explore our main proposition as they allow us to analyse user activity. In order to assess the width of segments targeted, we measure how Facebook users interact across multiple campaigns.

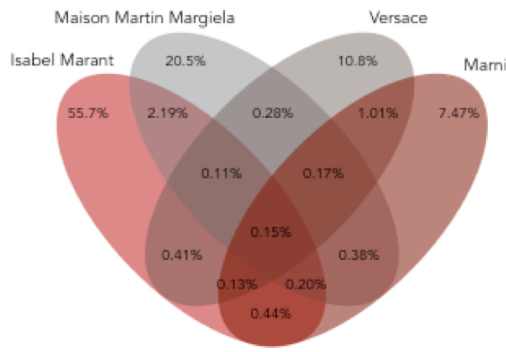
One noteworthy thing is that even though Beyoncé has a larger follower base on her own Facebook fan page (63m) than David Beckham (54m) she has fewer users who have acted on her H&M campaign than David Beckham. This might be because David Beckham's campaign appeals to both men and women, whereas the Beyoncé campaign mostly appeals to women. Figure 6(c) shows how users engage across H&M's designer collaborations. It should be noted that the overlaps exist, but surprisingly small. All overlaps only amount to 5.5%, meaning that only 5.5% of the users that have acted in one designer collaboration have also acted in at least in one another collaboration. The biggest overlap is between Isabel Marant and Maison Martin Margiela. The smallest overlap between just two collaborations is between Maison Martin Margiela and Versace, which are also two very different designer styles. To our surprise only 0.15% users have been active in all four campaigns. It might be is because each of these designers have a very different style.



(a) Style icons vs designers



(b) Style icon campaigns



(c) Designer collaborations

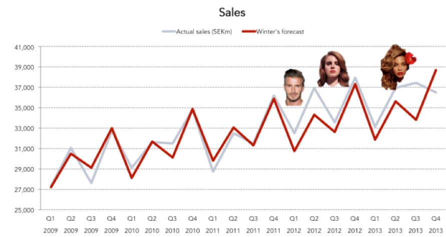
Figure 6: User engagement across different campaigns

To sum up for each Venn diagram we have:

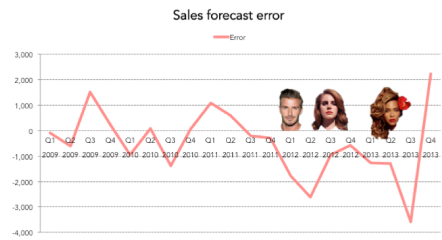
- Style icons and designer collaborations relatively small overlap $\implies$ P_4 *supported*
- Between style icons relatively small overlap $\implies$ P_4 *supported*
- Between designer collaborations relatively small overlap $\implies$ P_4 *not supported*

The overall overlap between designer collaborations (5.5%) is bigger than the overall overlap between style icons (3.6%), which is in line with P_4.

P_5-In a sales forecast, the quarters with style icon campaigns have the largest deviations between actual and forecasted sales: If style icons create more social buzz than designers, there should be bigger errors in the forecasted sales for quarters in which there is a style icon campaign. Assuming that social buzz is proportional and positively correlated with sales, if we observe bigger negative divergences (actual sales being larger than forecasted sales) in style icon quarters then it indicates that style icon do create more social buzz than designer collaborations. We have forecasted quarterly sales using the method presented in section 4.4 and figure 7(a) show actual quarterly sales and forecasted quarterly sales. The model used for forecasting 2012 and 2013 is: $F_{t+l} = [\ L+(t+l)T\ ] \times S_{t+l}$ with the values $L = 32,960, T = 316, S_1 = 0.91, S_2 = 1.02, S_3 = 0.96, S_4 = 1.09$. From figure 7(a) it can be seen that the



(a) Forecasted quarterly sales for Style Icon



(b) Forecast errors for Style Icons

Figure 7: Forecasted sales for Style Icon Campaigns

forecast follows historical sales (2009-2011) relatively accurately. From beginning of 2012 to end 2014 divergences between actual observed sales and forecasted sales become larger. As it can be seen on the graph the David Beckham campaign took place in Q1 2012, the Lana Del Rey campaign happened in Q3 2012 and the Beyoncé campaign in Q3 2013. Interestingly enough the actual sales in and around these periods are fairly bigger than forecasted to be. In figure 7 the forecast errors are plotted over time. P_5 is supported as the occurrence of the biggest forecast errors coincide with the style icon campaigns. Since there are no style icon campaigns during the historical sales data period this further improves the validity of our findings, because the effect of the style icon campaigns are reinforced.

6. Discussion and Conclusion

Empirical evaluation of the five sub-propositions revealed interesting and important facts about H&M style icon campaigns and designer collaborations that were not blindingly obvious. One could argue that since it was discovered in P_1 that style icon campaigns got more likes and fewer post than designer collaborations, the results from evaluating P_2, was downright obvious (style icon campaigns get more likes per post than designers collaborations). Furthermore, it is interesting in P_2 how large and rather consistent the difference is between likes per post for style icon campaigns and designer collaborations. Most of the empirical findings make good sense, for example style icons campaigns create more social buzz by getting the most comments and likes due to their significantly larger follower base on their own Facebook fan pages. However, it does not make sense why style icon campaigns do not also generate more posts than that of designer collaborations, which is quite an interesting for several reasons. It is contradictory with that assumption that style icons should create more social buzz than designer collaborations. It is also important because it shows that social buzz-wise; designer collaborations perform better than expected. It has not been possible to find a truly plausible reason for this, which makes the results even more interesting, though not very useful.

Further, it was fairly surprising that H&M posts make up a larger percentage of total posts for style icons than for designers collaborations, since it was expected that style icon campaigns would be more organically self-driven. However, it is worth noting that in absolute numbers, H&M does on average make more posts for designer collaborations. Evaluation of P_4 also disclosed some interesting and important findings. The overlap between designer collaborations was expected to have been a lot bigger than a total of 5.5% because it was expected that fashion enthusiast would be interested in all collaborations with large and famous designers. This could indicate that H&M's followers might have a more individualised taste and that H&M manages to reach a broader target than initially anticipated with their collaborations. One could notice that the total overlap between designer collaborations was bigger (5.5%) than the total overlap between style icons (3.6%). It was slightly surprising how relatively clear the relationship between the occurrence of style icon campaigns and large forecast errors was as per P_5. Prior to this it had been shown that style icons create more social buzz than designer collaborations, however, the main proposition is only supported if the underlying assumption about more social buzz equals more sales also proves to be true. The fact that the results from P_5 indicate that the underlying assumptions are not unrealistic makes the findings important.

Proposition	P_1	P_2	P_3	P_4	P_5
Empirical Finding	not supported	supported	not supported	not supported	supported

Table 5: Result of sub-propositions

The coverage of our five sub-propositions leaves us with our main proposition. As shown in table 5, there is no clear answer to our main proposition. The indicators of social buzz proved to be mixed, the aggressive social media marketing approach by H&M was not all that different between the two campaign types and also customer segments for all campaigns proved to be almost equally differentiated. The relationship between style icon campaigns and larger than expected sales was evident.

To summarize, data from Facebook was extracted and organized in specific campaigns. Visual and descriptive analytics based on Social Set Analysis was used, which revealed insights about the social buzz around the campaigns initiated by H&M. Further, financial data about quarterly sales of H&M was used in order to make a forecast of sales. Results show that H&M's style icon campaigns had a greater impact on sales than their designer collaborations. However, this is our best case interpretation of the data and we are aware that when working with big data there is a risk of claiming to objectivity and accuracy without further consideration, since data to some degree represent an *objective truth* [24]. Our conclusions and interpretation of data are based on the assumption of buzz marketing as well as the operationalisation of social buzz measured by amount of posts, comments and likes. If we had interpreted the data based on assumptions of brand loyalty, our findings could have been different. However, based on examination of related work and the theory of buzz marketing, we argue that the interpretation of our data and analytics is sufficient and provides adequate empirical grounds and warrants for the substantive conclusions.

6.1 Implications

Our findings support existing empirical work that shows a connection between big social data and business data [25]. Particularly, our findings show that there is a connection between social buzz and sales. We believe this is important, since it emphasizes the need for organizations to strategically engage in social media as well as for academic research to focus on investigating the connection between social buzz and organisational performance. Our findings have several direct and indirect implications for organizations (B2C). The direct implication of the results is the connection between social buzz on Facebook and the development in sales. The indirect implication is that organizations should strategically use Facebook as a means of creating social buzz, since this can, according to our analysis, positively impact the financial performance of the organization. Combined with prior knowledge of the damaging effect of negative reviews imply that creating social buzz using Facebook requires that organizations have domain specific knowledge of their industry and a thorough understanding of their customers in order to know what type of activities and campaigns will create a positive response and thereby a positive impact on sales. Overall, our findings highlight the value of using social media to achieve business goals.

6.2 Recommendations

It is recommended for H&M to continue running multiple campaigns simultaneously targeting different segments, to focus slightly more on style icon campaigns to create more social buzz, and push their own efforts for promoting their current and upcoming campaigns. Besides the above, our actionable insights are based on the belief that we have found empirical evidence of a positive correlation between social buzz and sales. We thus recommend H&M, to continue their allocation of resources to study and analyze multiple social media sites such as Facebook, Twitter, and Instagram, for them to get a better understanding as well as a more nuanced picture of their customers. By these means we believe H&M can better group customers and potential customers into segments and target campaigns towards them for a better effect.

6.3 Limitations

Our study is build around social media campaigns and their effect on our forecast. The limitation about campaigns is that we looked only at two years in isolation. In addition, we measure the effect of the campaigns based on Facebook posts, comments, and likes. Additional measures (such as other social media) might also have effects on campaigns, which we have not considered.

References

[1] National Research Council, *Frontiers in Massive Data Analysis*. The National Academies Press, 2013.

[2] V. Mayer-Schönberger and K. Cukier, *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt, 2013.

[3] E. Bakshy, M. Simmons, D. Huffaker, C.-Y. Cheng, and L. Adamic, "The social dynamics of economic activity in a virtual world," in *International AAAI Conference on Web and Social Media*, 2010.

[4] J. Bollen and H. Mao, "Twitter mood as a stock market predictor," *Computer*, vol. 44, no. 10, pp. 0091–94, 2011.

[5] D. H. Dorr and A. M. Denton, "Establishing relationships among patterns in stock market data," *Data & Knowledge Engineering*, vol. 68, no. 3, pp. 318–337, 2009.

[6] M. Gavrilov, D. Anguelov, P. Indyk, and R. Motwani, "Mining the stock market (extended abstract): which measure is best?" in *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2000, pp. 487–496.

[7] M. Kharratzadeh and M. Coates, "Weblog analysis for predicting correlations in stock price evolutions," in *International AAAI Conference on Web and Social Media*, 2012.

[8] M.-A. Mittermayer, "Forecasting intraday stock price trends with text mining techniques," in *System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on*. IEEE, 2004, pp. 10–pp.

[9] S. Asur and B. Huberman, "Predicting the future with social media," in *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on*, vol. 1, 2010, pp. 492–499.

[10] N. B. Lassen, R. Madsen, and R. Vatrapu, "Predicting iphone sales from iphone tweets," in *Enterprise Distributed Object Computing Conference (EDOC), 2014 IEEE 18th International*. IEEE, 2014, pp. 81–90.

[11] C. Seebach, I. Pahlke, and R. Beck, "Tracking the digital footprints of customers: How firms can improve their sensing abilities to achieve business agility," in *Proceedings of the 19th European Conference on Information Systems (ECIS); Helsinki, Finland*, 2011.

[12] P. vd Reijden and O. R. Koppius, "The value of online product buzz in sales forecasting." in *ICIS*. Citeseer, 2010, p. 171.

[13] T. Geva, G. Oestreicher-Singer, N. Efron, and Y. Shimshoni, "Do customers speak their minds? using forums and search for predicting sales," *ICIS 2013 Proceedings*, 2013.

[14] L. Wu and E. Brynjolfsson, "The future of prediction: how google searches foreshadow housing prices and quantities," *ICIS 2009 Proceedings*, p. 147, 2009.

[15] W. Zhang and R. Lau, "The design of a network-based model for business performance prediction," *ICIS 2013 Proceedings*, 2013.

[16] S. Nann, J. Krauss, and D. Schoder, "Predictive analytics on public data-the case of stock markets." in *ECIS*, 2013, p. 102.

[17] C. Oh and O. Sheng, "Investigating predictive power of stock micro blog sentiment in forecasting future stock price directional movement." in *ICIS*, D. F. Galletta and T.-P. Liang, Eds. Association for Information Systems, 2011.

[18] D. L. Hoffman and M. Fodor, "Can you measure the roi of your social media marketing?" *MIT Sloan Management Review*, vol. 52, no. 1, p. 41, 2010.

[19] J. . Andzulis, N. G. Panagopoulos, and A. Rapp, "A review of social media and implications for the sales process," *Journal of Personal Selling & Sales Management*, vol. 32, no. 3, pp. 305–316, 2012.

[20] M. J. Hatch and M. Schultz, "Of bricks and brands:: From corporate to enterprise branding," *Organizational Dynamics*, vol. 38, no. 2, pp. 117–130, 2009.

[21] S. Chopra and P. Meindl, *Supply chain management. Strategy, planning & operation*. Springer, 2007.

[22] A. Hussain and R. Vatrapu, "Social data analytics tool: Design, development, and demonstrative case studies," in *Enterprise Distributed Object Computing Conference Workshops and Demonstrations (EDOCW), 2014 IEEE 18th International*, Sept 2014, pp. 414–417.

[23] A. Hussain, R. Vatrapu, D. Hardt, and Z. Jaffari, "Social data analytics tool: A demonstrative case study of methodology and software." in *Analysing Social Media Data and Web Networks*, M. C. Rachel Gibson and S. Ward, Eds. Palgrave Macmillan, 2014 (in press).

[24] D. Boyd and K. Crawford, "Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon," *Information, communication & society*, vol. 15, no. 5, pp. 662–679, 2012.

[25] R. Vatrapu, R. R. Mukkamala, A. Hussain, and B. Flesch, "Social set analysis: A set theoretical approach to big data analytics," *IEEE Access*, vol. 4, pp. 2542–2571, 2016.