Forecast Adidas' sales using Google Trends Data Containing Different Key Words Representing Brand Image

Jiepeng Lian (jili16ab) *Kin Kwan Chen (kich16ac)*

***Abstract*-This paper test whether Google Trends data and Facebook data can provide an accurate forecast on Adidas quarterly global sales. The effect of brand image towards sales is also assessed by including Google Trends data of several key words representing brand images. The Facebook dataset consists of Adidas' Facebook page with the number of likes, comments, posts etc. Financial data in terms of quarterly global sales figures are collected from Adidas' financial report. The finding suggest that Google trends data does have values in predicting sales. Regressions models including seasonal dummy variables predict Adidas' quarterly sales with a high accuracy. Finding of this paper shows that Google trends data and social media data can provide valuable information for retailing companies for their planning and supply chain management.**

Ⅰ - Introduction

In this paper, we show how big social data can be used to analysis brand coolness as well as predict financial outcome like sales of Adidas AG (Adidas).

The rise of social media provides a platform for easy interaction between people as well as freely expressing their emotion. For companies that are selling goods to customers, social media can serve as a new channel of hearing customer's feeling and feedback of their products. Using the AIDA-framework, it can be deducted that attention from social media may lead to purchases, which enable data analysts to predict future sales. Research show that the large volume of data from social media can be used to predict sales of iPhones (Lassen, Madsen, & Vatrapu1, 2014) with a high accuracy.

Brand image is a heavily used term by marketing practitioners. It is an important concept in consumer behavior research since the early 1950s (Goldberg, Gorn, & Provo, 1990). It is correlated with uniqueness, brand conscious and many other aspects (Rahman & Cherrier, 2010). Research confirms that brand image like brand coolness have positive correlation towards brand effect, brand trust and brand loyalty, and therefore customers are more willing to pay a premium for the cooler brand (Hodis, 2010). We will analyze what kinds of brand image influences sales in this paper.

In this paper, we chose Adidas AG (Adidas) as the target company. Adidas AG is a German company that designs and manufactures foot wears, clothing, apparel and accessories. It is the largest sportswear manufacturer in Europe and the second biggest in the world, ranked 90th in Forbes Most Valuable Brands. Adidas can also be regarded as a fashion company since its light fashion brand Adidas Originals is very active in social media, gain much attention in Facebook, Twitter and Instagram. Adidas Originals was Instagram's most liked sneaker account of 2015. Therefore, we can get a vast amount of data from Adidas social media. As Adidas is listed in German, its quarter financial report which is including sales and revenue can be found and used for investigation. Therefore, Adidas is a great company to analyze the relationship the relationship between sales and social media data. Adidas had a poor performance in 2013 and 2014, but Adidas manage to turn their business around in 2015 and 2016. One key element that contributes their regaining glory is increasing brand image of coolness. Social media has played a significant role in Adidas' rising brand coolness. So Adidas also serves really well to analyze the how brand image effect sales.

Research Questions：

1) How and to what extent Facebook data can predict Adidas quarterly sales?
2) How and to what extent Google data can predict Adidas quarterly sales?
3) How different brand images key words predict Adidas' quarterly sales?
4) Is there any difference between people's behaviors in social media and searching website?

Ⅱ- Related works

Accurate sales forecast is one of the key success factors to apparel retails companies (S. Thomassey, 2010). Many international retails companies has a very long global supply chain and it is significant for them to order several months in advance to ensure that enough amount of goods will be sent in certain period.

Sales are dependent to many variables, and data from social media are now found relevant to this important indicator towards business (Rapp, 2012). The modern forecast method applies sophisticated calculation algorithms on any number of macroeconomic and microeconomic variables, including weekly sales data, prior year results, preseason and trended product-level plans (Bauer, 2015). Researchers are now able to forecast global sales of IPhone with data from Twitter in a very high accuracy (Lassen, Madsen, & Vatrapu1, 2014). Sales of another international apparel retailer H&M has also been found to be forecasted by Facebook pages' data (Rosenborg, et al., 2016).

Brand images has long be regarded as a key success factor of retailing. Research argue that the brand image has a direct effect towards sales volume (Ataman, 1992). However, it is difficult to find one clear, objective definition of brand image. One significant kind of brand image "coolness" can serve as a good example. There is a large number of synonyms of cool: awesome, dope, sweet etc. (Bergh & Behrer, 2016). The definition of different brand images differ widely among people, and it is common that a company that shows on the list of 'the coolest brands' is regarded by some people as not cool at all.

Google Trends data is freely available data based on Google searches that shows how frequently a word or phrase is searched for by region or globally, and Google trends data are now found to be very valuable for forecasting future performance. Choi and Varian have used Google Trends data to forecast near-term values of economic indicators, like sales (Choi & Varian, 2012). Their research shows that forecast models with Google trends data have very high accuracy, surpassing other models by 5% to 20%. Google trends data are also found correlation with traffic in physical stores and online stores performance like traffic of physical stores (Bauer, 2015).

Ⅲ- Methodology

A. Case Company Description

Adidas has been seen as the top 2 sportswear brands for a very long time. Adidas performed not very well and experienced a 2-year decline since 2014, many of their business decision have been proved to be unprofitable. One typical case is the acquisition of Reebok. In 2015, Adidas was surpassed by Under Armour in the U.S. sportswear market, giving up its long time second place in the market of highest volume. However, they managed to score a comeback since late 2015 by hiring new CEO of North America area as well as applying a business strategy of focusing more on footwear (Kell, 2016). Recently Adidas brand images became increasingly cool and fashion, so we believed that it would serve as a good target company for our research questions.

B. Dataset Description

1. Google Trends Data

Google Trends shows how often a particular searching-term is entered relative to the total searching volume. The index is based on the total volume of searching for one term divided by the total number of queries for the same period. It also allows the users to compare the search volume among several terms. We decided to take Nike as a benchmark to quantify the brand image of Adidas by Google Trends. Nike Inc. (Nike) is also a global sportswear manufacturing company, ranked 18th most valuable brands by Forbes. The reason we choose Nike as a standard of coolness is that Nike has long been regarded as one of the coolest brands in the world like Apple. Nike has a very cool and fashion brand image and Nike is in the exact same industry of Adidas, so it is reasonable to compare this two rivals. We choose terms "cool", "Teens", "Awesome" and "Star" as the indicators of brand image, combine them with terms "Adidas" or "Nike", then compare their search volumes. We prefer using data from searching engine, because Google Trends data is less polished comparing with social media data, and therefore practice better in the analysis of customer's opinion. Searching for a brand along with cool words represent the apprehension of Adidas and relevant to Adidas' brand image.

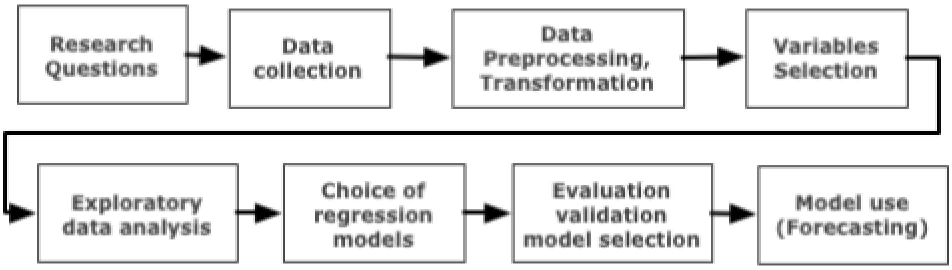
2. Facebook Data.

We mainly focus on Adidas Facebook page. Adidas Facebook page has 25,321,237 page likes at December 4, 2016. We use Social Data Analytics Tool (SODATO) to collect data from Adidas page. We can get access to data of many attributes of Facebook pages like numbers of likes, comments and shares. However, due to the limitation of the software, we can only analyze a two-year data from Facebook, which is too short for a valuable analysis. So we decide to use the Facebook data analyze to compare with the Google data analyze and find out the difference between these two data, trying to figure out how different people behave on searching website and social media.

3. **Financial Data**

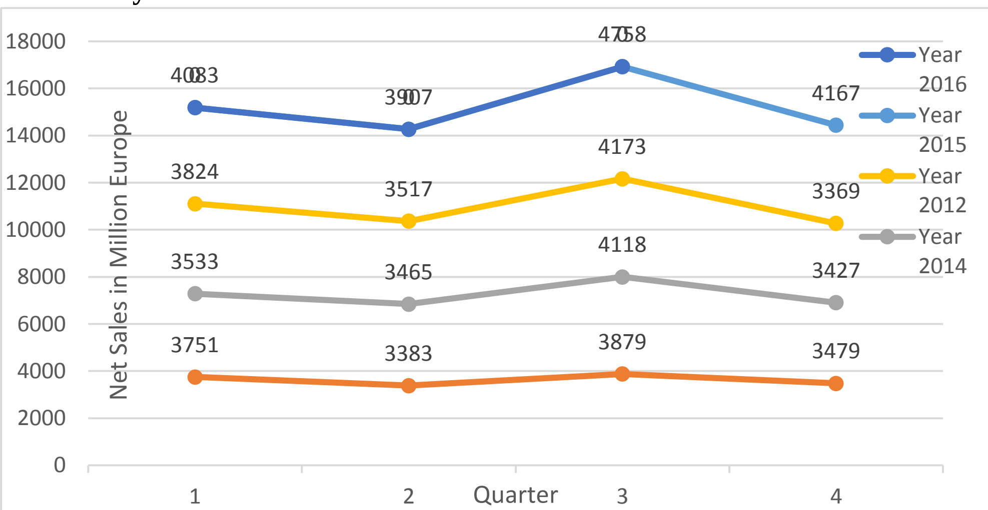
Adidas is listed in Xetra so the financial data need to be exposed to public quarterly. The financial data including sales of Adidas can be easily collected in Adidas' quarterly financial reports.

C. Data Analytics Process



We followed these steps for the forecasting analysis. We use exploratory analysis performed to investigate the relationships between the dependent and independent variables. For each regression, we examine the coefficients as well as the adjusted R-square to consider to what extent the regressions models fit the data we collect. Based on the statistical measures, the models that explain variation in the sales data best are used for forecasting. Using most suitable regression models, we forecast the global net sales of Adidas in the first three quarters of 2016. The forecasts are then compared to the actual global net sales on the quarterly financial report of Adidas to examine the accuracy of the models.

D. Data Analysis



After visualizing the Google trends data, we find that there are clear seasonal patterns in Adidas' net sales.

In order to remove this seasonal effect on sales, the first step is deseasonalizing the raw Google data. We use two different methods which have different implications to deseasonalize.

In the first method, we calculate the season weight by dividing quarter sales by the total sales of that year, then divide this season weight we get by 0.25 to get the seasonal influence coefficient. We multiply all the independent variables by the coefficient of the year before regression. This implies that the seasonal effects are proportional to the sales.

The second method is adding season dummy variables *dummy1 dummy2* and *dummy3* as independent variables into regression. *Dummy1* has a value of 1 if the data is in the first quarter, otherwise the value is 0. *Dummy2* and *dummy3* are similarly defined. If the data is in the fourth quarter, all three dummy variables are 0. This method implies that the seasonal effect on sales is directly increasing or decreasing sales in a certain amount.

Another important issue is the time lag effect. Customers may be attracted by sports stars by an epic game and search for more information about them, where they find Adidas new launched shoes that are endorsed by the stars. However, the purchasing of the equipment will take some time since customers purchase the endorsed product after becoming the fans of the stars, which require continues following. As there can be a time lag between the search on Google and the actual purchasing, lag of each explanatory variable except dummy variables are including in the regression. Since the financial data are collected quarterly, we set time lag in the unit of the quarter. We do a regression for each searching terms with a maximum of three-quarter lag. Every lag variable with less lag time than maximum lag time will be included in the regression. For example, in the three-quarter lag regression, one-quarter lag explanatory variable L.x and two-quarter explanatory variable L2.x are also included.

Ⅳ- Result

A. Searching-term "Adidas"

Firstly we analyze the Google data of searching-term "Adidas" to get an overall picture of how accurate Google data can predict the global sales. We use STATA for regression and we get the p-value, R-square and adjusted R-square of each regression.

Method	**Maximum Lag setting**	**p-value**	**R-square**	**Adj. R-square**
Method 1	no time lag	0.0002	0.6409	0.6153
	1Q lag	**0.0001**	**0.8022**	**0.7693**
	2Q lag	0.0009	0.7959	0.7346
	3Q lag	0.0040	0.8248	0.7373
method 2	no time lag	**0.0000**	**0.8980**	**0.8610**
	1Q lag	**0.0001**	**0.9192**	**0.8743**
	2Q lag	0.0011	0.9275	0.8654

	3Q lag	0.0046	0.9538	0.8891

We assess these regression base on the adjusted R-square. The higher the adjusted R-square, the better the model explain the historical data. We pick the model with the highest adjusted R-square for each method. In both method the model with a maximum one-quarter lag has the highest adjusted R-square. These two models are then used to test the forecast accuracy on the testing sample. The predicted net sales can be calculated by setting the exact value of the dependent variable in each model.

Forecast Model	predicted sales	actual sales	diff %
Method1 with Q1 time lag	4588.683	4769	-3.93%
	4105.171	4422	-7.72%
	5165.158	5413	-4.80%
Method2 with Q1 time lag	4775.649	4769	0.14%
	4643.374	4422	4.77%
	5306.980	5413	-2.00%

B. Searching-term "Adidas Cool"

We then analyze word "Adidas" plus term "Cool"

Method	**Maximum Lag setting**	**p-value**	**R-square**	**Adj. R-square**
Method 1	no time lag	0.1372	0.1508	0.0901
	1Q lag	0.2439	0.2096	0.0778
	2Q lag	0.2057	0.3542	0.1605
	3Q lag	0.0515	0.6546	0.4819
method 2	no time lag	0.0915	0.4894	0.3038
	1Q lag	0.1104	0.5810	0.3482
	2Q lag	0.2697	0.5818	0.2234
	3Q lag	0.2135	0.7474	0.3939

We find that all the models have a high p-value and low adjusted R-square, which means that the relationship between explanatory variables and dependent variables are not significant, and the model explain the low percentage of the fluctuation of past sales data. Therefore, we find it unnecessary to test the models.

C. Searching-term “Adidas Star”

D.

Method	Maximum Lag setting	p-value	R-square	Adj. R-square
Method 1	no time lag	0.0003	0.6215	0.5944
	1Q lag	**0.0000**	**0.8759**	**0.8552**
	2Q lag	0.0001	0.8837	0.8488
	3Q lag	0.0006	0.8915	0.8372
method 2	**no time lag**	**0.0000**	**0.9304**	**0.9051**
	1Q lag	**0.0000**	**0.9454**	**0.9150**
	2Q lag	0.0004	0.9464	0.9005
	3Q lag	0.0053	0.9511	0.8827

Similar to the case of searching-term “Adidas”, the one-quarter lag model has the highest adjusted R-square in both methods. The testing results are:

Forecast Model	predicted sales	actual sales	diff %
Method1 with Q1 time lag	4051.833	4769	-17.70%
	3796.367	4422	-16.48%
	4552.604	5413	-18.90%
Method2 with Q1 time lag	4339.595	4769	-9.90%
	4026.999	4422	-9.81%
	4654.747	5413	-16.29%

E. Searching-term “Adidas Awesome”

Method	Maximum Lag setting	p-value	R-square	Adj. R-square
Method 1	no time lag	0.0013	0.5339	0.5007
	1Q lag	0.0039	0.6025	0.5363
	2Q lag	0.0074	0.6832	0.5882
	3Q lag	**0.0042**	**0.8231**	**0.7347**
method 2	no time lag	0.0358	0.5792	0.4262
	1Q lag	0.0848	0.6093	0.3922
	2Q lag	0.1033	0.7046	0.4514
	3Q lag	0.0642	0.8571	0.6571

We take the three-quarter time lag model using method1 as the testing model. We decide not taking any model from method2 since they all have high p-value (higher than 0.05), implying

that in the model using seasonal dummy variables, the relationship between global net sales and searching-term “Adidas Awesome ” is not significant.
The testing results are:

Forecast Model	**predicted sales**	**actual sales**	**diff %**
	4312.673	4769	-10.58%
	3792.180	4422	-16.61%
Method1 with Q3 time lag	4338.244	5413	-24.77%

F. Searching-term “Adidas Teens”

G.

Method	**Maximum Lag setting**	**p-value**	**R-square**	**Adj. R-square**
	no time lag	0.0006	0.5780	0.5478
	1Q lag	0.0002	0.7542	0.7132
	2Q lag	0.0004	0.8235	0.7706
Method 1	**3Q lag**	**0.0009**	**0.8798**	**0.8197**
	no time lag	**0.0004**	**0.8260**	**0.7627**
	1Q lag	0.0028	0.8312	0.7374
	2Q lag	0.0094	0.8630	0.7457
method 2	3Q lag	0.0292	0.8991	0.7578

We select the three-quarter time lag model using method1 and the no time lag model using method2 to test the accuracy.

Forecast Model	**predicted sales**	**actual sales**	**diff %**
	4561.588	4769	-4.55%
	4924.395	4422	10.20%
Method1 with Q3 time lag	5656.553	5413	4.31%
	4838.126	4769	1.43%
	4708.827	4422	6.09%
Method2 with Q1 time lag	5336.206	5413	-1.44%

Ⅴ- Discussion

A. About Google Trends Data

Basically Google Trends data serves as a good tool to forecast net sales. Google Trends data on "Adidas" is able to provide a forecast of the global sales in a relatively high accuracy. Using the one-quarter model of method2 we can forecast sales in 2016 with difference less than 5%. We can also develop models that have adjusted R-square higher that 75% in other searching-terms. In sum, Google Trends data shows a really good potential to forecast global net sales.

B. About Method 1

Using Method 1 to deseasonalize, one-quarter time lag models have the highest adjusted R-square in searching terms "Adidas" and "Adidas Star", while three-quarter lag models have the highest adjusted R-square in searching terms "Adidas Awesome" and "Adidas Teens". Looking deep into the regression models, we find an interesting trait in the coefficients: the regression coefficients of the one-quarter lag variable (L.x) are always negative in all four regressions. In some models, the p-values of the one-lag explanatory variable are low (less that 0.05) which indicate that the negative effect on sales is statistically significant. This implies that keeping other explanatory variables remained, the increase of last quarter's searching queries will actually decrease the sales forecast of this quarter. One possible explanation of decreasing sales is that people search on the Internet about Adidas and they find criticisms on Adidas. The image of a brand consists of the way it is perceived by consumers (Zimmer & Golden, 1988). Once customers have bad impressions on certain products, they may share their negative comment with their friends, making them deciding not to buy Adidas' products and therefore decrease the sales. Interpersonal communications have long been recognized as an influential source of information for consumers and negative words can make a great difference to sales (Sonnier, Leigh, & Rutz, 2011). The negative effects of criticism of sales may have lag because people can poorly comment the products even if after they see the trailers as they find any disappointed features. The effect of praise comments, on the contrary, would be more instant on the sales, since people are more likely to give positive comment after the see or try the actual products.

For the longer lag effect, two-quarter time lag and three-quarter time lag, have positive coefficients in the highest adjusted R-square model. This may imply that positive effect of comments sustain a long period, while negative effects may only influence short-term sales. However, the p-values of the coefficients of longer lag explanatory variable vary a lot with different searching-terms. In term "Adidas Awesome" the p-values are 0.022 and 0.028, while they are 0.056 and 0.464 in terms "Adidas Teens". "Awesome" is a complete pleasant word while "Teens" is a neutral word, the p-values indicate that praising words have statistically significant positive effects on sales. Searching activity two or three quarters ago might have high explanatory power due to clothing or apparel items that are not purchased as frequently as light fashion clothes and foot gears.

C. *About Method 2*

Method 2 assumes that seasonal patterns would be absolute sales numbers rather than a proportion of sales. We find that using method 2, short time lag model have better performances. For searching-terms “Adidas” and “Adidas Stars”, the one-quarter time lag model perform better using method 2 to deseasonalize, and both adjusted R-squares are very high (0.8743 and 0.9150). It seems that method 2 deseasonalize better than method 1. However, the issues of collinearity appear in this method especially in the models have long lag times. Different from Model 11 which divided the seasonal weighted before regression, method 2 include seasonal dummy variables in the regression. As a result, we include lag effects and seasonal effects in regression at the same time, and lag explanatory variables and seasonal dummy variables in the regressions may have a certain level of correlations, which may cause collinearity issues. In order to explicitly divide seasonal effects and lag effects, more research is needed to be done. In these models the lag effect of the explanatory variables are relatively small, and therefore the collinearity issues are less serious than models with long lag times. However, we find that the issues still exist and they greatly affect the statistical significance of the variables. After including only a one-quarter lag in the regression of term “Adidas”, the p-value of no lag explanatory variable increase from 0.000 to 0.323. Similar change happens in the regression of term “Adidas Star”. It seems that seasonal effects represented by the coefficients of seasonal dummy variables do have collinearity issues with lag effects. We also find that in these two cases, the p-values of dummy variable *dummy2* are high (0.764 and 0.563), which means that there is no statistical difference between the sales in the second quarter and in the fourth quarter. We decide to simply remove lag explanatory variables as well as dummy variable *dummy2* for in order to improve the regression models of all searching-terms. The results are:

Explanatory Variables	**p-value**	**R-square**	**Adj. R-square**
"Adidas"	0.0000	0.8972	0.8715
"Adidas Cool"	0.0395	0.4879	0.3599
"Adidas Star"	0.0000	0.9281	0.9102
"Adidas Awesome"	0.0131	0.5785	0.4731
"Adidas Teens"	0.0002	0.8013	0.7516

After polishing the model, we figure out 3 terms with impressive adjusted R-square and low p-values: “Adidas”, “Adidas Star” and “Adidas Teens”. We then use the testing sample to test the accuracy of these three models:

Explanatory Variables	predicted sales	actual sales	diff %
	4679.008	4769	-1.92%
	4389.388	4422	-0.74%
"Adidas"	5132.309	5413	-5.47%
	4576.596	4769	-4.20%
	4242.206	4422	-4.24%
"Adidas Star"	4867.135	5413	-11.22%
	4606.139	4769	-3.54%
	4377.651	4422	-1.01%
"Adidas teens"	5055.195	5413	-7.08%

Basically the accuracy is high in the two recent quarters and it decreases in the third coming quarters.

Further theoretical analysis and calculations are needed to be done in order to divide lag effects and seasonal effects and build a solid theoretical basis for the regression models.

D. *About Searching Words*

We find that while words “Star” and “Teens” are combined with “Adidas”, linear regression models can explain a large percentage of historical data and provide an accurate forecast on global net sales. However, the performance turns down when some words representing positive meaning are combined, like “Cool” and “Awesome”. Linear regression functions better with neutral words. People who search for these words have a relatively constant probability to buy Adidas’ products, as they are likely to hold no or slightly personal feeling in it. People search Adidas along with praising words are already holding certain positive opinions on Adidas’ products. These kinds of people are more likely to buy them, and the probability change based on how strong their perceptions on the products, which cannot be measured by searching queries. So the relationship between positive words and global net sales is more complicated, and linear regression models are not able to explain this relationship.

E. About “Star”

Among the 5 searching-terms, our deseasonalize regression models perform the best with data from searching-term “Adidas Star”. However, the forecast accuracy of the regression in sales of 2016 are not as impressive as the high percentage of historical data the regression model can explain. The forecasts of the models are always less than the actual net sales. One possible reason is that not only people search for “Adidas Star” more frequently on Google, but also people are more likely to buy Adidas' product after the searching. Therefore, stars endorsement are increasingly critical in building brand image and pushing Adidas' net sales.

F. About Facebook data.

The overall Facebook data perform poorly in explaining the Adidas net sales. Most of the regressions models have large p-value and small, sometimes negative adjusted R-square.

explanatory variable	**p-value**	**R-square**	**Adj. R-square**
TotalPosts	0.5187	0.0726	-0.082
TotalPosts lag Q1	0.0858	0.477	0.3724
TotalPosts lag Q2	0.549	0.0965	-0.1294
TotalPosts lag Q3	**0.0472**	**0.7796**	**0.7062**
TotalLikes	0.2347	0.2253	0.0962
TotalLikes lag Q1	0.4164	0.1356	-0.0373
TotalLikes lag Q2	0.1873	0.3869	0.2336
TotalLikes lag Q3	0.9187	0.0041	-0.3279
TotalComments	0.2201	0.2379	0.1109
TotalComments lag Q1	0.4008	0.1442	-0.0269
TotalComments lag Q2	**0.105**	**0.5217**	**0.4021**
TotalComments lag Q3	0.962	0.0009	-0.3321
TotalCommentLikes	0.5771	0.0547	-0.1028
TotalCommentLikes lag Q1	0.5002	0.0954	-0.0855
TotalCommentLikes lag Q2	**0.1483**	**0.4442**	**0.3052**
TotalCommentLikes lag Q3	0.9114	0.0048	-0.3269
TotalSharers	0.6836	0.0296	-0.1321
TotalSharers Lag Q1	0.7	0.0323	-0.1613
TotalSharers Lag Q2	0.2117	0.3554	0.1942
TotalSharers Lag Q3	0.9887	0.0001	-0.3332

The three-quarter time lag total posts regressions function relatively good, with the lowest p-value and highest adjusted R-square. Compare with the Google data that one-quarter time lag models have the best performance, Facebook data has a longer time lag to influence the net sales. Facebook is a platform for create and sharing information, people create Facebook data about Adidas when they see new post or advertisement of new products on Adidas Facebook pages, it takes some time for them to know the products or certain event and decide to purchase or not. As for Google data, customers may have the initial idea of purchasing before using Google and they look for more information like other's comments and opinions to make the last decision. The harder they feel in deciding, the more they will use Google. So large percentage of Google data are created short periods before final purchasing. Therefore, Google data has a shorter lag time on net sales than Facebook data.

Ⅵ- Conclusion & Recommendation

Facebook data is not able to predict Adidas quarterly global net sales with a considerable accuracy. However, Google trends data can predict quarterly sales within error of 10%. Analysis on Google Trends data also shows that objective words that link to brand image like “Stars” can predict Adidas sales well, while words “cool” cannot forecast the sales. Google data predict the short future (0 to 1 quarter lag) sales well, while Facebook data performs better in forecasting the long-term (2 to 3 quarters lag) future.

Firstly we recommend Adidas to keep on signing contracts with big name stars and sports teams. Signing stars can significantly build up Adidas’ brand image and we found that people are increasingly interested in sports of fashion stars as the searching frequency climbed up in the past 5 years. People are also more likely to buy products that link to those celebrities.

We also recommend Adidas to continue studying and analyzing multiple social media sites such as Facebook, Twitter, and Instagram. These social media data can provide Adidas a better understanding of their customers and what influence them the most. Adidas Original is the Instagram account that has the most followers among the sportswear industry. We believe studying about these data will enable Adidas to better group customers and target them by effective promotion and therefore is able to build up brand coolness as well as increase sales.

Ⅶ- Limitation

We only choose a few words that correlated to brand image base on our own expectation. However, these considerations may be wrong, some other significant word terms may provide a better insight of the relationship between brand image and global sales.

The deseasonalize method we use in this paper is based on two different assuming seasonal patterns. However, the actual seasonal effect on the sales may be more complicated. Further research on deseasonalization of retail sales needs to be done in order to clarify how seasonal patterns influence the sales.

We are only able to access limited Facebook pages within a very short period, so we do not pay much effort in analyzing Facebook data. There are probability much gold in Facebook data. Another significant social media that we are unable to access data is Instagram. Adidas’ Instagram followers increase dramatically in 2015 and Instagram played an important role in Adidas’ comeback in that year.

References

Ataman, B. (1992). A note on the effect of brand image on sales. *Journal of Product & Brand Management*.

Bauer, S. (2015). *Using Google Trends to Predict Retail Sales*. Retrieved from pwc: http://www.pwc.com/us/en/retail-consumer/publications/assets/pwc-using-google-trends-to-predict-retail-sales.pdf

Bergh, J. V., & Behrer, M. (2016). *How Cool Brands Stay Hot: Branding to Generations Y and Z.* London.

Choi, H., & Varian, H. (2012). Predicting the present with google trends. *Economic Record Vol 88*, 2-9.

Goldberg, M. E., Gorn, G., & Provo, R. W. (1990). In Search of Brand Image: a Foundation Analysis. *Advances in Consumer Research Volume 17*, 110-119.

Hodis, M. A. (2010). WHY IS APPLE COOL? AN EXAMINATION OF BRAND COOLNESS. *2010 AMA Educators' Proceedings Volume 21*, 147-148.

Kell, J. (2016, May 4). *Here's Why Adidas Is Scoring a Comeback in the U.S.* Retrieved from Fortune: http://fortune.com/2016/05/04/adidas-scores-a-comeback/

Lassen, N. B., Madsen, R., & Vatrapu1, R. (2014). Predicting iPhone Sales from iPhone Tweets. *Enterprise Distributed Object Computing Conference*, (pp. 81-90).

Rahman, k., & Cherrier, H. (2010). Correlates of Cool Identity: Humor, Need For Uniqueness, Materialism, Status. *Advances in Consumer Research Volume 37,*, 886-888.

Rapp, A. (2012). A Review of Social Media and Implications for the Sales Process. *Journal of Personal Selling & Sales Management Vol.32*, 305-316.

Rosenborg, D. C., Buhl-Andersen, I., Nilsson, L. B., Rebild, M. P., Mukkamala, R. R., Hussain, A., & Vatrapu, R. (2016). *Buzz vs. Sales: Big Social Data Analytics of Style.* Copenhagen.

S. Thomassey. (2010). Sales forecasts in clothing industry: The key. *International Journal of Production Economics vol.128*, 470-483.

Sonnier, G. P., Leigh, M., & Rutz, O. J. (2011). A Dynamic Model of the Effect of Online Communications on Firm Sales. *Marketing Science*, 702 - 716.

Zimmer, M. R., & Golden, L. L. (1988). Impressions of Retail Stores: A Content Analysis of cosumer Images . *Journal of Retailing*, 265.