

Mapping the Sensitivity of the Public Emotion to the Movement of Stock Market Value: A Case Study of Manhattan

Yuhao. Kang^a, Jialin. Wang^a, Yueyao. Wang^a, Stephan Angsuesser^a, Teng. Fei^{a,*}

^a School of Resource and Environmental Sciences, Wuhan University, China – feiteng@whu.edu.cn

Commission IV, WG IV/3

KEY WORDS: Affective computing, Economy, Manhattan, Stock market, Sensitivity map, Face detection, Cognition recognition

ABSTRACT:

We examined whether emotion expressed by users in social media can be influenced by stock market index or can predict the fluctuation of the stock market index. We collected the emotion data by using face detection technology and emotion cognition services for photos uploaded to Flickr. Each face's emotion was described in 8 dimensions the location was also recorded. An emotion score index was defined based on the combination of all 8 dimensions of emotion calculated by principal component analysis. The correlation coefficients between the stock market values and emotion scores are significant ($R > 0.59$ with $p < 0.01$). Using Granger Causality analysis for cause and effect detection, we found that users' emotion is influenced by stock market value change. A multiple linear regression model was established ($R\text{-square}=0.76$) to explore the potential factors that influence the emotion score. Finally, a sensitivity map was created to show sensitive areas where human emotion is easily affected by the stock market changes. We concluded that in Manhattan region: (1) there is an obvious relationship between human emotion and stock market fluctuation; (2) emotion change follows the movements of the stock market; (3) the Times Square and Broadway Theatre are the most sensitive regions in terms of public emotional reaction to the economy represented by stock value.

* Corresponding author

1. INTRODUCTION

Recently, with the development of the social sensing movement, there is a tendency to use social media data to understand users' behaviours (Liu et al. 2015). Thanks to the spread of location-based techniques and wireless devices, hundreds of thousands of data are recorded with highly spatio-temporal resolution, which is a new and valuable dataset for analysing human's behaviour and urban development. With spatial analysis, data with location information extracted from social media offers rich insights of the society (Yang, Mu and Shen 2015). One of the important properties that can be extracted is human emotion.

Human cognition and emotions are innate and can provide meaningful aspects of human analysis (Perikos and Hatzilygeroudis 2016). Human emotion can be influenced by numerous factors, including genes, cultural attitudes and human behaviours as well as natural or social environments such as weather condition or economy conditions etc. (CAPR 2011, Ryu 2008, Koots 2011). Previous researches of behaviours economics (Smith 2003), behavioural finance (Nofsinger 2005) and socioeconomics hypothesis have also critically examined that human emotion can affect individual behaviour such as investment behaviour (Parker 2007). That means economy can be driven by massive investors with certain emotion. Thus, it seems that the economical dynamic and the emotional condition are interrelated and interact as both cause and effect. Under this circumstance, it is worth to explore if public emotions have any significant relationship with the fluctuation of stock market values. If so, does the market value influence or is it influenced by emotions? What is the time lag if there is a correlation between the two?

Some researches tried to explore the relationship between the economy and the public emotion expressed by social media (Li 2014). Machine learning has been used as a powerful tool to extract emotional information from social networks. To name a few, Bollen et al. (2011) used Twitter mood to predict the stock market and improve the accuracy of the Dow Jones Industrial Average predictions. Nofer and Hinz (2015) used tweets from Twitter to predict German stock market and the results showed that it is necessary to take the spread emotion into account for predicting stock market value. Cohen (2013) collected emotion words from the newspaper to predict stock market value and the finding shows that pleasant moods predict increasing prices and vice versa. While Li et al. (2014) who also used news articles for predicting Hong Kong Stock Exchange prices found that a sentiment model cannot provide useful predictions. Strau et al. (2016) investigated emotions in Dutch newspapers and their effects on stock market value. They found that newspapers reflected movements on the stock market value the following days. Whether the results were successful or not, whether emotion change lag behind stock market value or not, some relationships exist between stock market value and emotions. However, researches mentioned above all extracted and qualified emotion for affective computing based on text. Tweets, news reports are the most popular data sources for extracting emotion, but they have some drawbacks and limitations.

Firstly, people send messages and write articles to express emotion after the events. That means, emotion conditions recorded are not in real-time and will be modified and transferred after a period of time. The buffer period may cause users to be dispassionate and to calm down so that moods expressed will become second-hand data. Secondly, it is difficult to classify and qualify the degree of the emotion condition accurately. For example, how to take "happy" and

"very happy" into consideration? How much happy is "very happy" more than "happy"? Lacking of hierarchy and highly accurate evaluation, no standard method can be utilized to illustrate the slight difference. Another challenge is the cultural universality. Different cultures, regions, nations have different customs, languages, expressing habits. There is no standard model to resolve and translate the differences among different cultures so far.

Considered that emotions can be expressed not only by speech and writing forms, but also by tones of the voice, facial expressions, body gestures (Chakraborty et al. 2009, Hudson et al. 2015), we focused on some other data sources. Though each culture has its own verbal expression, people from all over the world have similar facial expressions for representing emotions. Researches have proved that facial expression are universal in some aspects when expressing basic emotions, which can be traced back to our primate ancestors (Wilhelm et al. 2014, Price. 2016). And now, with the advances in computer vision, machine learning, and artificial intelligence, it is possible to extract emotion expressions from photos with faces (Ali Khan et al. 2015). Compared with text based emotion extracting process, emotions obtained from images with faces have three advantages: (1) facial expressions are recorded in real-time when emotions are revealed spontaneously; (2) emotions detected by cognitive services and algorithms will be represented with an accurate confidence and score, which can be easily quantified for further computing and analysis. (3) for problems being multi-lingual, it can overcome the cross-cultural difference because of its higher universality.

In addition, though several researches extracted emotions from enormous social media data for further analysis, most of them focused on time series of emotion fluctuations. For instance, Yang et al. (2015) studied the effect of climate and seasonality of the depression mood from Twitters' users in the U.S. Few researches explored the geographic patterns of emotions (Yang and Mu 2015), let alone the geographic patterns of emotional sensitivity. What can influence the distribution of emotions? And where are the most sensitive regions? Are there some regions' where certain emotions are more sensitive than in other regions?

With questions and hypothesis put forward above, in this paper, we attempt to explore the following: (1) providing a new method for extracting emotion information accurately from facial expressions in photos users uploaded; (2) analysing whether public emotions could be influenced by or could predict the stock market value; (3) locating the area with highest correlation between emotion and stock market value, and the areas with more sensitivity than others; (4) analysing what factors might influence the relationship between emotion and stock market.

2. METHODOLOGY

2.1 Data preparation

As one of the most influential financial centre in the world, there are several important stock exchanges and financial firms located in Manhattan, New York City including: Dow Jones & Company located at the Wall Street who publish Dow Jones Industrial Average (DJIA), Standard & Poor's (S&P), which is located at Water Street, and the National Association of Securities Dealers Automated Quotation (NASDAQ) located at the Broadway Street. On the one hand, according to the first law of geography, near things are more related than distant thing

(Tobler 1970). Hence, in order to exploit the potential relationship between stock market value and emotions, Manhattan was selected as our study area. On the other hand, considering that numerous visitors flow into Manhattan every day, there is a higher possibility to get adequate photo sources than other sites. We collected the same time series of these 3 stock market indexes from Yahoo Finance. In total, 60 months of index values were harvested, and the average and standard deviation index values of the stock market were calculated for a monthly time series.

Then, Flickr photos were used as our data source for further research. Flickr is a public image hosting website where users can share personal photographs. With time and geographical information recorded automatically by GPS and other equipment embedded in smartphone, more and more photos are uploaded to Flickr with geo-tagged information. Till December 2014, more than 92 million photos were uploaded to the album from users all over the world. In this research, photos uploaded from January 1st, 2012 to December 31th, 2016 in Manhattan were collected with their geo-attributes created by users using Flickr API (Flickr, 2015). Each photo's record contains its unique photo id, user id, location information (latitude and longitude), original url and the uploading date and time. After collecting data from Flickr, a heat map was created to visualize the photo distribution in Manhattan.

2.2 Affective computing

Only those photos with faces can be utilized for further analysis. Most of photos collected from Flickr are scenery without faces and should be discarded first. We applied the Face++ platform, which is a publicly available vision technology website for face recognition to determine which photos have faces (Face++. 2015). It can detect and locate human faces within an image and returns high accuracy face bounding boxes. All photos stored were posted to this application and those with faces marked.

After selecting the photos with faces, we used Microsoft Cognitive Services, the Emotion API to detect each face's emotion (Microsoft 2015). By uploading photos as input, it returns the confidence of each face with a set of emotions, including eight different dimensions of human mood, namely *anger*, *contempt*, *disgust*, *fear*, *happiness*, *neutral*, *sadness*, and *surprise*. The value of confidence of each emotion ranges from 0 to 1.

2.3 Principal component analysis and emotion score

Not all dimensions of mood have equal importance for exploring the relationship between emotion and stock market index. Thus, Principal Component Analysis (PCA) was used to reduce the dimensions. PCA is the most popular and classical unsupervised linear technique for dimensionality reduction, with unified Emotion Score (ES) as one output to express the emotion condition for further analysis (Zhong and Enke 2017).

Before performing PCA, there is necessity to normalize all dimensions of emotion confidences and stock market values to z-scores on the basis of their means and standard deviations by Eq. 1:

$$Z = \frac{X - \mu}{\sigma} \quad (1)$$

where Z is the normalized value, X is the emotion value, μ is the mean of all emotion values and σ is the standard deviation.

To achieve this goal, a variance-covariance matrix of a linear of the data X , $var(X)$ is needed. Let λ_i be the eigenvalues of the correlation matrix $var(X)$ such that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M$. Then, let the vectors $e_i^T = (e_{i1} \ e_{i2} \ \dots \ e_{iM})$ denote the eigenvalues of $var(X)$ corresponding to the eigenvalues λ_i , $i=1, 2, \dots, M$. It turns out that the elements of these eigenvalues are the coefficients of the principal components. Thus, the principal components of the standard data can be written as Eq. 2:

$$Y_i = \sum_{j=1}^M e_{ij} X_j, \quad i = 1, 2, \dots, M \quad (2)$$

where Y_i is one component. In addition, to enable the comparison of stock market value and emotion score, it is necessary to choose principle components that can explain the data best and also simplify the data structure.

After the process of PCA, we defined an index of emotion score, which combines and compresses all dimensions of emotions for further analysis:

$$E_s = a_1 X_1 + a_2 X_2 + a_3 X_3 + a_4 X_4 + a_5 X_5 + a_6 X_6 + a_7 X_7 + a_8 X_8 \quad (3)$$

where E_s is the emotion score of one face, X_1 is the confidence value of anger, X_2 is that of contempt, X_3 is that of disgust, X_4 is that of fear, X_5 is that of happiness, X_6 is that of neutral, X_7 is that of sadness, and X_8 is that of surprise.

At last, we computed each correlation coefficient between the emotion score and the stock market indexes to find whether there is some relationship between the affection and stock.

2.4 Correlation analysis

After computing the exact emotion score of each face, we considered the question whether stock market values and emotion scores are correlated. Hence, the correlation value between stock market values and emotion scores following Eq. 4 were computed:

$$\text{corr}(X, Y) = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \quad (4)$$

where E is the expected value operator, X and Y are two vectors, μ is the mean of all values and σ is the standard deviation.

In order to visualize the correlation pattern, line graphs of each index were created. From the graphs, it is easy to recognize the tendency of fluctuations.

2.5 Granger causality analysis

Results of the correlation show that stock market and emotions are related to each other. The next step is to investigate that what is the cause and what is the effect between public emotion and stock market value. In other words, whether stock market value could influence or be predicted by public emotion. To answer this question, we applied Granger Causality Analysis, which aims at determining whether one time series is useful in forecasting another (Granger 1969). It assumes that if a time series of X is said to Granger-cause Y , then changes of values of Y will exhibit a statistically significant lagged than that of X . We then calculated two indexes for the analysis followed Eq. 5 and Eq. 6:

(a) Monthly change of the emotion score D_{es} :

$$D_{es} = E_s(m_1) - E_s(m_2) \quad (5)$$

(b) Monthly change of the average stock market value D_{value} :

$$D_{value} = Value(m_1) - Value(m_2) \quad (6)$$

where m_1 is the target month and m_2 is the month before, E_s is the emotion score and $Value$ is the stock market index. Eviews software was used to investigate whether the stock market values influence emotion scores. Average monthly stock market values of 3 stock indexes were obtained as input with monthly averaged emotion score.

2.6 Multiple linear regression

Multiple linear regression, which aims at figuring out the relationship between several independent or predictor variables and a dependent or criterion variable, was then employed to explore more details about which stock market factors can influence the emotion score. We extracted and computed six independent features from stock market indexes and emotion scores, including: (1) monthly averaged stock market value; (2) and (3): fluctuation of the emotion score of the last 2 months; (4) and (5): fluctuation of the stock market of the last 2 months; (6) monthly count of photos. The fluctuation mentioned above means the difference between the value of month X_1 and the value of month X_2 followed the Eq. 7:

$$F = X_1 - X_2 \quad (7)$$

where F is the fluctuation value, and the X_2 is the month prior to the month X_1 .

In order to make sure the six features we extracted are independent, which means two or more predictor variables in a multiple regression model should not be highly correlated to avoid the multicollinearity problem, we used several methods to test the degree of multicollinearity. O'Brien (O'Brien 2007) suggested to use a formal detection-tolerance or the variance inflation factor (VIF) for multicollinearity according to the Eq. 8 and Eq. 9.

$$\text{tolerance} = 1 - R_j^2 \quad (8)$$

$$\text{VIF} = \frac{1}{\text{tolerance}} \quad (9)$$

where R_j^2 is the coefficient of determination, which is a number that represent the proportion of the variance in the j th independent variable that is associated with the other independent variables in the model. *Tolerance* values of less than 0.20 or 0.10, and a VIF of more than 10 conforms that there is a serious multicollinearity problem. In addition, collinearity diagnostics were also analysed, which included another two evaluation indexes: eigenvalues and condition numbers. If eigenvalues are close to 0 or condition numbers are above 30, the predictors are highly intercorrelated and the regression may be significantly multicollinearity (Atkinson 1984, Fildes 1993).

2.7 Explore the sensitivity region

No researches before detected the geographical sensitive pattern of emotion distribution. Sensitivity region means areas where emotion conditions have high correlation to the stock market value. In order to explore which areas are sensitive, research areas are separated into different blocks based on latitude and longitude. The latitude of our research area ranges from 40.6894° N to 40.8322°N and the longitude ranges from 73.9175°W to 74.0334°W. The map was separated into a 15×15 matrix of grids. In each block, the correlation coefficient between emotion

scores of faces in it and stock market value change were calculated with their p-value. To locate the highest sensitive region, we created a map to render the results of correlation coefficients. Blocks whose correlation coefficient ranges from 0 to 0.001 will be discarded because that means there is no adequate data.

3. RESULTS

3.1 Data descriptions

Three stock market indexes, DJIA, NASDAQ, and S&P, were obtained from Yahoo Finance. To conveniently compare them, a line graph of the three indexes was created. Each index has been normalized using the Z-score method so that values are in the range of -2 to 2 in general. From Figure 1 it can be concluded that the three indexes are similar, all reflecting some aspects of the U.S economy.

Using Flickr API, 820,438 photos were obtained within the Manhattan Island and then stored in a cloud database. Figure 2 shows the distribution of photos based on kernel density estimation. According to the heat map, photos are densely clustered in red regions, including Broadway Theatre, Times Square, Wall Street and south of the Central Park.

Following the methodology explained in section 2.2, we utilized Face++ API and Microsoft Cognitive Services for affective computing. After that, 127,982 faces or 15.6% of the photos remained for further processing.

3.2 Emotion Score

PCs	Component proportion	Cumulative proportion
1	0.9179	0.9179
2	0.0497	0.9676
3	0.0171	0.9847
4	0.0100	0.9947
5	0.0033	0.998
6	0.0011	0.9991
7	0.0001	0.9999
8	0	1

Table 1 The results of PCA over the entire data

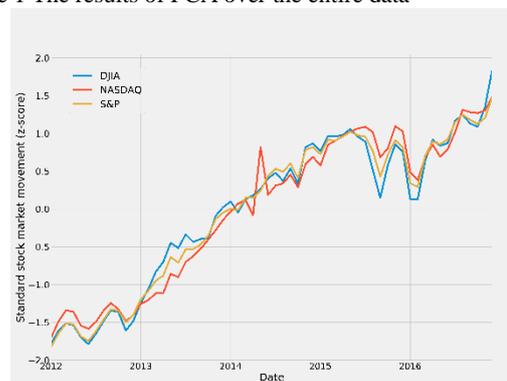


Figure 1 Values for three stock market indexes

Using PCA, the contribution rate of 8 principal components of the data can be computed. These principal components based on PCA for the entire data are ordered by their importance and are shown in Table 1. It can be inferred that the first principal component listed in Table 1 can explain more than 90% (91.79%) variation of the data set.

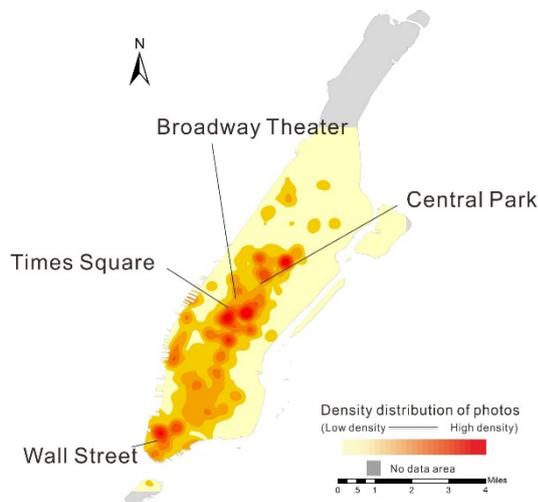


Figure 2 Density distribution of photos at Manhattan

Hence, according to the cumulative proportion and the eigenvector, emotion score was defined by the following Eq. 3'.

$$E_s = 0.0099X_1 + 0.0061X_2 + 0.0026X_3 + 0.0016X_4 - 0.7361X_5 + 0.6761X_6 + 0.0247X_7 + 0.015X_8 \quad (3')$$

Based on the coefficients of X_i in Eq.3', we can infer that happiness and neutral are two dominant key emotions of the final emotion score. Note that happiness has negative influence on the emotion score while neutral has positive influence. As this, seems to contradict our common sense, we redefined the emotion score as its negative, that is:

$$E_{sn} = -E_s \quad (10)$$

where E_{sn} is the new and final emotion score used in the later analysis. Correlation coefficient between the emotion score and happiness was calculated and visualized in order to illustrate the relationship of the two.

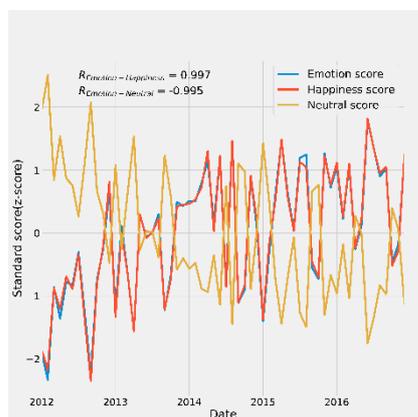


Figure 3 Correlation between happiness score / neutral and emotion score

Figure 3 shows that the correlation coefficient between happiness and emotion score is 0.997, and -0.995 between neutral and emotion score. That means happiness and neutral are almost opposite in this research.

3.3 Correlations between stock market values and emotion scores

The following 3 diagrams (Figure 4) show the correlation between emotion scores and 3 stock market values. Results show that all of them have high correlation values (about 0.6) with high significance (p -value < 0.01). Intuitively, the tendency of emotion scores in the graphs is similar to the lines of the three stock market indexes. All of them increased in the past 5 years, and the results proved that is true. In detail, it can be observed that there are slight differences of correlation coefficients among the three indexes (no more than 0.015). S&P has the highest correlation coefficient with emotion score (0.609) while NASDAQ has the lowest correlation coefficient (0.594).

3.4 Granger causality analysis results

Table 2 shows the result of Granger causality analysis. Based on the result, we can reject the null hypothesis that the emotion score can predict the stock market value with a high level of significance. We observed that the emotion score had the highest Granger causality relation with stock market when lag was 1. With lags increasing, F-statistic decreased and the p -value increased. Thus, stock market fluctuation influences the emotion conditions, but is not predictable.

3.5 Multiple linear regression results

Six features were taken into consideration: (1) fluctuation value of target month's average emotion score and last month's average emotion score; (2) fluctuation value of last month's average emotion score and the month before last month's average emotion score; (3) monthly average stock market value; (4) fluctuation value of target month's average stock market value and last month's average stock market value; (5) fluctuation value of last month's average stock market value and the month before last month's average stock market value; (6) the amount of photos.

(c)

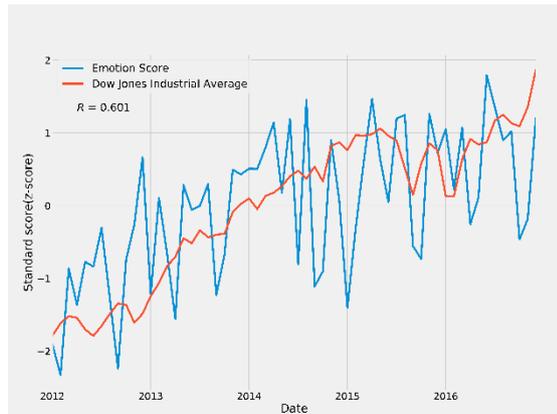
Figure 4 shows the results of multiple regression analysis. It can be observed that models with different stock market values have similar results. R-square floated from 0.761 (NASDAQ) to 0.766 (DJIA, S&P) with high significance of p -value < 0.01 .

Table 3 shows the standardized regression coefficients of multiple regression models. The interesting results show that coefficients of emotion scores fluctuation and stock market values have great influence on emotion scores and are significant. Models with different stock markets values have similar results. With the standardized coefficient, it is easy to answer the question which of the independent variables have a greater effect on the dependent variable. The absolute values of the coefficients show the weight of emotion scores. Among three variables which are significant, emotion score fluctuation between values and last month's values is the most important feature that influence stock market value and has great significance.

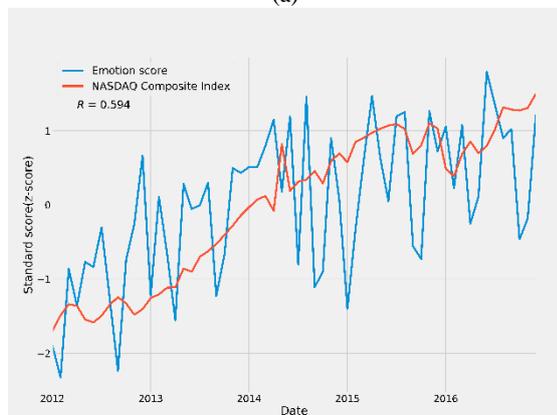
Stock market values also have great effect and significance. DJIA has the greatest effect of 0.5708 while NASDAQ has the lowest effect of 0.5676. Emotion score fluctuations of the past 2 months are also related to the emotion score and range from 0.3409 (with DJIA) to 0.3471 (with S&P).

In order to test the reliability of the model and the multicollinearity problem, we followed the process mentioned in section 2.6. Table 4 shows the results of these diagnosis

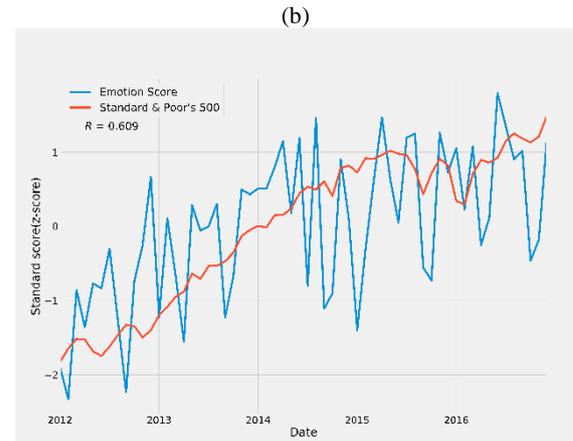
indexes. It can be observed that the tolerance values are more than 0.20, and VIF values are less than 10. No eigenvalues are close to 10 and condition index values are all less than 30. All these statistics show that there is less multicollinearity problem and prove that our models are reliable and can be used for further analysis.



(a)



(b)



(c)

Figure 4 Line graph of correlation between emotion scores and (a) DJIA values; (b) NASDAQ values; (c) S&P values

3.6 Visualization of sensitive regions

The sensitivity map of emotion score to the stock market was designed to illustrate where the sensitive regions are. From Figure 6, it can be observed that Broadway Theatre, Times Square, which are located in Midtown of Manhattan have a higher correlation coefficient (0.622 and 0.410).

Penn Station also has a high correlation level of 0.511. Statistics demonstrated that results in these regions are significant. The region where the World Trade Centre and the Wall Street are located also have a positive correlation of the stock market with 0.234.

Null Hypothesis:	Lag (Month)					
	1		2		3	
	F-Statistic	Prob.	F-Statistic	Prob.	F-Statistic	Prob.
Emotion_score does not Granger Cause DJIA	1.02871	0.3148	0.50179	0.6083	3.03976	0.0374
DJIA does not Granger Cause Emotion_score	16.4380	0.0002**	7.33649	0.0015**	3.58567	0.02*
Emotion_score does not Granger Cause NASDAQ	0.48264	0.4901	0.21459	0.8076	2.77248	0.051
NASDAQ does not Granger Cause Emotion_score	14.4327	0.0004**	6.66509	0.0026**	3.27817	0.0284*
Emotion_score does not Granger Cause S&P	0.97381	0.328	0.49222	0.614	3.11365	0.0344
S&P does not Granger Cause Emotion_score	15.8570	0.0002**	7.05581	0.0019**	3.44181	0.0236*

* $p < 0.05$

** $p < 0.01$

Table 2 Granger Causality Analysis Result

	Coefficient	p-value	Tolerance	VIF
Const	0.0604	0.377		

Emotion score fluctuation 1	0.6368	0**	0.747	1.338
Emotion score fluctuation 2	0.3474	0**	0.801	1.249
DJIA	0.5702	0**	0.973	1.028
Stock market fluctuation 1	-0.7006	0.027*	0.937	1.067
Stock market fluctuation 2	-0.4811	0.137	0.943	1.061
Amount of faces	-0.0198	0.804	0.907	1.103

Table 3(a) Standardized regression coefficients of multiple regression models (DJIA)

	coefficient	p-value	Tolerance	VIF
Const	0.0151	0.827		
Emotion score fluctuation 1	0.6204	0**	0.760	1.315
Emotion score fluctuation 2	0.3411	0**	0.802	1.247
NASDAQ	0.5671	0**	0.981	1.019
Stock market fluctuation 1	-0.4390	0.176	0.885	1.130
Stock market fluctuation 2	-0.0094	0.977	0.891	1.122
Amount of faces	0.0159	0.842	0.935	1.069

Table 3(b) Standardized regression coefficients of multiple regression models (NASDAQ)

	coefficient	p-value	Tolerance	VIF
Const	0.0435	0.537		
Emotion score fluctuation 1	0.6322	0**	0.746	1.340
Emotion score fluctuation 2	0.3484	0**	0.793	1.261
S & P	0.5655	0**	0.994	1.006
Stock market fluctuation 1	-0.5873	0.159	0.949	1.054
Stock market fluctuation 2	-0.4112	0.326	0.960	1.042
Amount of faces	-0.0008	0.938	0.930	1.075

Table 3(c) Standardized regression coefficients of multiple regression models (S & P)

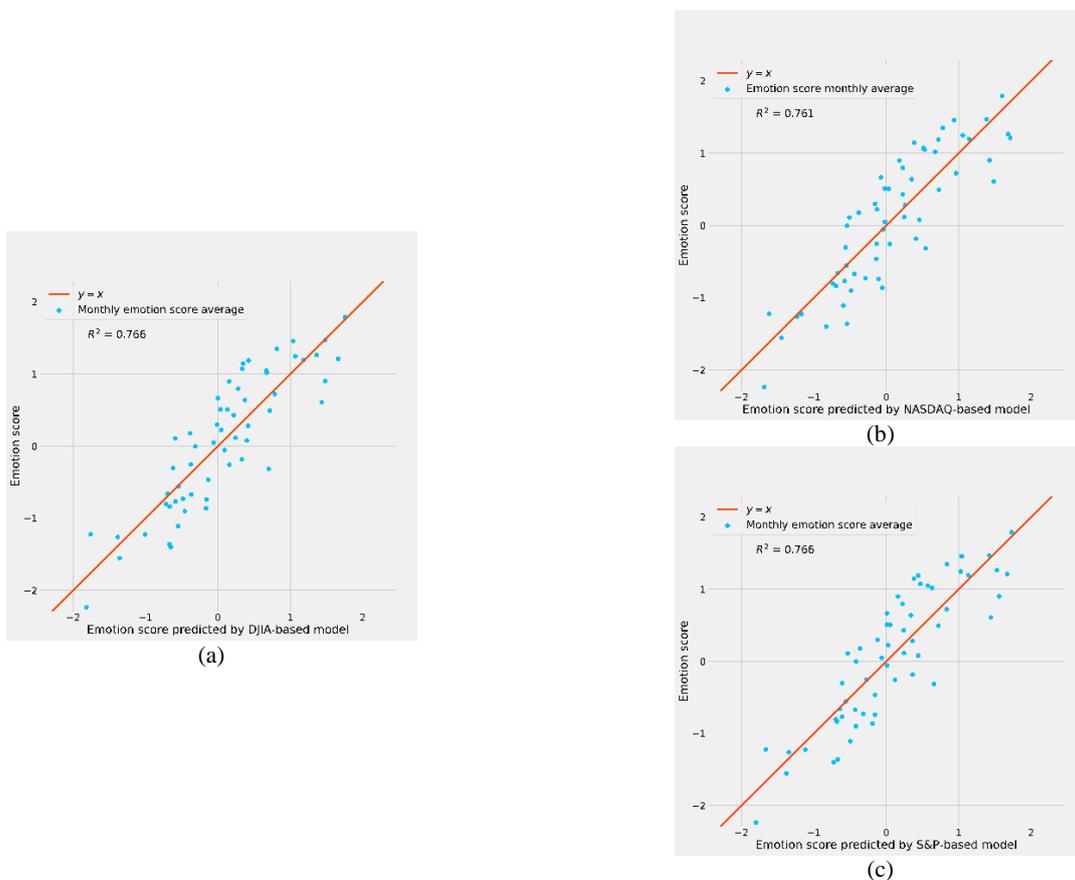


Figure 5 Multiple regression results for emotion scores

Dimension	Eigenvalue	Condition index
1	1.594	1
2	1.436	1.054
3	1.096	1.206
4	0.923	1.314
5	0.795	1.416
6	0.701	1.508
7	0.455	1.872

Table 4 (a) Eigenvalue and condition index of DJIA

Dimension	Eigenvalue	Condition index
1	1.547	1
2	1.328	1.079
3	1.174	1.148
4	1.026	1.228
5	0.863	1.339
6	0.613	1.588
7	0.449	1.857

Table 4 (b) Eigenvalue and condition index of NASDAQ

Dimension	Eigenvalue	Condition index
1	1.638	1
2	1.470	1.055
3	1.061	1.243
4	0.940	1.320
5	0.775	1.454
6	0.661	1.574
7	0.456	1.896

Table 4 (c) Eigenvalue and condition index of S & P

4. CONCLUSION AND DISCUSSION

Intrigued by the facts and discussions that emotions and economy are interrelated, we investigated whether public emotions extracted and computed from facial expressions in photos uploaded to Flickr are correlated to the stock market movements, and where the most sensitive regions in terms of emotions related to the economy are. Results show that among the 8 dimensions of emotion, happiness and neutral are the two key dominant moods. Human emotions in Manhattan respond to the stock market fluctuation and movements with significant correlation coefficients and lag behind it. That means emotion changes will follow the stock market movements. Multiple linear regression results show that stock market values have a great and highly significant effect on the emotion score and the results are reliable. Based on the correlation coefficient, a map representing the emotion sensitivity to the stock market was created. The map shows that the Times Square and Broadway Theatre are regions most sensitive to the stock market movements.

Also, there are several important factors that are not considered in this research but will be examined in the future. First of all, we hope that our analysis is not limited to this particular geographical location and crowd. As Flickr's users usually upload their trip photos, data we collected are mostly from tourists who may not live there. Whose emotions are better to reflect the stock market values? Local people or people not living there? We should examine users' information in detail and explore the potential explanations.

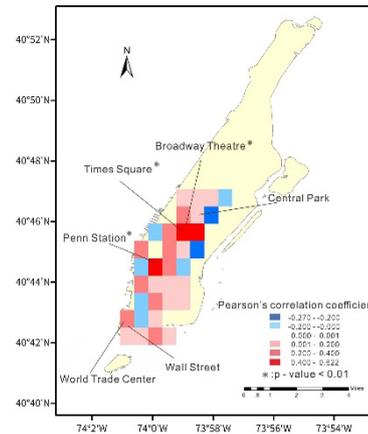


Figure 6 Sensitive map of emotion score to the stock market

Secondly, in this research, the locations where most photos were accumulated happened to be the most sensitive regions. We will further analyse to what extent the sensitivity relies on the sampling size and to what extent it relies on the location. Last but not least, more deep and detailed analysis is necessary to explore the essential reasons to find how stock market movements could cause fluctuations in human emotions. Besides Manhattan Island, there are several other financial centres in the world. Do they have the same phenomenon? Further analysis will take them into consideration as well to find out if there are general patterns and explanations.

ACKNOWLEDGEMENTS

We want to thank Yu Liu in Peking University, for his impressive suggestions of our methods and insights. And Stephan Anguesser, who helps us improving our writing.

REFERENCES

- Ali Khan, S., N. Riaz, S. Akram & S. Latif (2015) Facial expression recognition using computationally intelligent techniques. *Journal of Intelligent & Fuzzy Systems*, 28, 2881-2887.
- Atkinson, A. C. (1984) Review: Two Books on Regression Diagnostics; D. A. Belsley, E. Kuh, R. E. Welsch, *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*; R. D. Cook, Sanford Weisberg, *Residuals and Influence in Regression*. *Annals of Statistics*, 12, 392-401.
- Bollen, J., H. Mao & X. Zeng (2011) Twitter mood predicts the stock market. *Journal of Computational Science*, 2, 1-8.
- Chakraborty, A., A. Konar, U. K. Chakraborty & A. Chatterjee (2009) Emotion Recognition From Facial Expressions and Its Control Using Fuzzy Logic. *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans*, 39, 726-743.
- Cohen-Charash, Y. C.-M. J. B. M. (2013) Mood and the Market: Can Press Reports of Investors' Mood Predict Stock Prices? *PLOS ONE*, 8, -15.
- Face++. (2015) Face++ API <https://www.flickr.com/services/api/>.

- Fildes, R. (1993) Conditioning diagnostics - collinearity and weak data in regression *Journal of the Operational Research Society*, 44, 88-89.
- Flickr. (2015) Flickr API <https://www.flickr.com/services/api/>.
- Granger, C. W. J. 1969. *Investigating Causal Relations by Econometric Models and Cross-spectral Methods*.
- Hudson, S., M. S. Roth, T. J. Madden & R. Hudson (2015) The effects of social media on emotions, brand relationship quality, and word of mouth: An empirical study of music festival attendees. *Tourism Management*, 47, 68-76.
- Institute of Medicine (US) Committee on Advancing Pain Research, C., and Education. 2011. *Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research*. Washington (DC): National Academies Press (US).
- Koots, L., Aallik, J. 2011. *The Influence of the Weather on Affective Experience An Experience Sampling Study*.
- Li, Q. T. P. L. Q. Y. Z. (2014) The effect of news and public mood on stock movements. *INFORMATION SCIENCES*, 278, 826-841.
- Li, X., H. Xie, L. Chen, J. Wang & X. Deng (2014) News impact on stock price return via sentiment analysis. *Knowledge-Based Systems*, 69, 14-23.
- Liu, Y., X. Liu, S. Gao, L. Gong, C. Kang, Y. Zhi, G. Chi & L. Shi (2015) Social Sensing: A New Approach to Understanding Our Socioeconomic Environments. *Annals of the Association of American Geographers*, 105, 512-530.
- Microsoft (2015) Emotion API <https://azure.microsoft.com/en-us/services/cognitive-services/emotion/>.
- Nofer, M. & O. Hinz (2015) Using Twitter to Predict the Stock Market. *Business & Information Systems Engineering*, 57, 229-242.
- Nofsinger, J. R. (2005) Social Mood and Financial Economics. *Journal of Behavioral Finance*, 6, 144-160.
- O'Brien, R. M. (2007) A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41, 673-690.
- Parker, W. D. (2007) The Financial/Economic Dichotomy in Social Behavioral Dynamics: The Socioeconomic Perspective. *Journal of Behavioral Finance*, 8, 84-108.
- Perikos, I. & I. Hatzilygeroudis (2016) Recognizing emotions in text using ensemble of classifiers. *Engineering Applications of Artificial Intelligence*, 51, 191-201.
- Price., M. (2016) Facial expressions-including fear-may not be as universal as we thought. *Science*.
- Ryu, K., S. 2008. *Influence of restaurant's physical environments on emotion and behavioral intention*.
- Smith, V. L. (2003) Constructivist and ecological rationality in economics. *American Economic Review*, 93, 465-508.
- Strau, amp & N. n. s. u. n. R. V. Piet (2016) Lagging behind? Emotions in newspaper articles and stock market prices in the Netherlands. *Public Relations Review*, 42, 548-556.
- Tobler, W. R. (1970) COMPUTER MOVIE SIMULATING URBAN GROWTH IN DETROIT REGION. *Economic Geography*, 46, 234-240.
- Wilhelm, O., A. Hildebrandt, K. Manske, A. Schacht & W. Sommer (2014) Test battery for measuring the perception and recognition of facial expressions of emotion. *Frontiers in Psychology*, 5, -23.
- Yang, W. & L. Mu (2015) GIS analysis of depression among Twitter users. *Applied Geography*, 60, 217-223.
- Yang, W., L. Mu & Y. Shen (2015) Effect of climate and seasonality on depressed mood among twitter users. *Applied Geography*, 63, 184-191.
- Zhong, X. & D. Enke (2017) Forecasting daily stock market return using dimensionality reduction. *Expert Systems with Applications*, 67, 126-139.