

Paper 499-2013

Impact of London Olympics According to Tweeters

Yu Fu, Oklahoma State University; Shirmeen Virji, Oklahoma State University; Goutam Chakraborty, Oklahoma State University

ABSTRACT

The theme of our paper is the impact of London Olympics 2012 on the economy of England as stated by people on social media. Overspending money on Olympics by the host country with the hope of giving a huge boost to the economy is an age old phenomenon. With an estimated cost of £9.3 billion (\$14.8 billion), the economy of England is expected to get a boost, according to the British government, as there were 300,000 people in London to attend the games on top of 600,000 tourists. Only time will tell the real economic impact, but for now we are more interested in knowing what people have to say about the spending on the Olympics. For the purpose of analyzing sentiments of the people, we collected tweets between July 20 and August 15 by putting a filter of "Olympics AND economy". We have used Text Mining node on SAS Enterprise Miner 7.1 and SAS Sentiment Analysis Studio to build the models and analyze the results. Our results showed us that negative sentiments prevail among people and their confidence in the British economy is minimal. Even though the government of England is very optimistic about the boost that they anticipate in near future, people are certainly pessimistic. We get text topics characterized by words such as 'failure' and 'empty' to describe the economy. We further divided the data into negative tweets and positive tweets to compare the word 'Olympic' by building concept linking diagrams. In the positive cluster, we get words such as good, boost, and growth etc. In the negative cluster, we get words such as empty, city, and Greece. The dataset is also run on Sentiment Analysis Studio to corroborate our findings. Towards the end, we have also created a trend that tracks changes in the sentiments of people over the three-week period during which the Olympics were going on.

INTRODUCTION

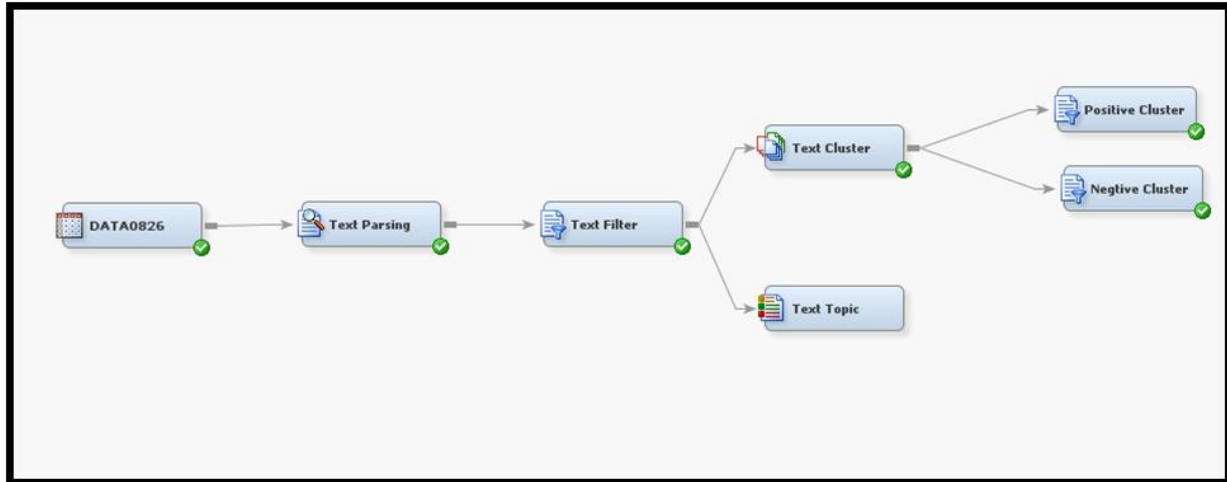
According to Oxford's Said Business School, every host country has overspent by an average of 179% since 1960. Atlanta, in 1996, spent 147% over original estimates, whereas Montreal, in 1976, went 796% over budget. Greece spent 60% over its set budget when it hosted Olympics in 2004 and, because of this, went through their current financial turmoil. It is a general observation that Olympic games almost always get their host countries into financial troubles. Bent Flyvbjerg, co-author of University of Oxford's research about Olympics, blames cost overruns on two things: optimism and politics. In this paper we will only emphasize on the optimistic part and analyze if people are also as optimistic as the British government about the most expensive Olympics in the history, at \$14.8 billion. In this digital world, where people express their sentiments on social media such as Twitter, Facebook, blogs, discussion forums etc., it is not that difficult to find out about the reactions of the people regarding the government spending on the Olympic Games. This paper only focuses on the sentiments that people expressed on Twitter and whether the tone of their tweets is negative, positive or neutral. We would also like to know if there is any change in the sentiments over three weeks of time.

METHODS USED

We used two methods to perform analysis on the data that was collected. The analysis is done by using SAS Text Miner and SAS Sentiment Analysis Studio.

TEXT MINING

In order to collect the data, we used topsy.com to gather the historical data of Twitter (Larson 2010). We collected the data from July 20th 2012 to August 12th 2012 by using the combination of keywords such as "Olympics+economy" and "Olympics+economics". We used Regular Expression rules to clean the data for URLs, RT tags and all the unnecessary special characters (Garla and Chakraborty, 2011). We used Text Miner of SAS® Enterprise Miner TM 7.1 to perform the analysis. We first parsed and filtered the data and then used Text Cluster and Text Topic nodes to analyze our data. Five clusters and various interesting topics were generated by these nodes. Two concept linking diagrams were used to find out the positive and negative sentiments of the people about the keywords "Economy" and "Britain." Furthermore, we divided our tweets into two sections: positive tweets and negative tweets. We created two more concept linking diagram on the same word, Olympics.



Display 1: Application of text mining on Twitter data

SENTIMENT ANALYSIS

In order to enhance our previous findings on the results of SAS Enterprise Miner, we gathered approximately 1,000 tweets between August 13th and August 15th from Twitter API. We used that data for testing purposes to corroborate our results. We used SAS Sentiment Analysis Studio 1.4.0 to validate our results by analyzing the sentiments of the people regarding the impact of the Olympic Games on the economy of England. We selected 139 positive and 192 negative tweets that represented the data that we used in text mining to create the corpus to train the model. We divided our data into 80% training and 20% validation and used the Basic Statistics model to get the final results. The probability threshold was set at 50% to determine whether the document is positive or negative.

RESULTS

USING SAS ENTERPRISE MINER

Text Topics

The Text Topics that are formed in SAS Enterprise Miner are very interesting. Some of the topics have pessimistic tone such as recession, ghost town, money, spend etc., whereas some of the topics have optimistic tone such as investment, UK GDP, billion, money.

Topic	# Docs ▼
+recession,economy,+economic,people,+greece	218
+uk,+economy,+bring,+recession,gdp	207
+london,+economic,+town,+summer,+ghost	206
+london2012,+global,+global economy,+win,+gold	148
+money,+spend,+billion,people,+bring	143
+host,+city,+country,+host,+london	141
economy,+look,+investment,+economy stutter,+stutter	139

Display 2: Text Topics created

Clusters Analysis

Five clusters were created by the Text Miner on the dataset. From our findings, we believe that two clusters (1 & 2) are positive; two clusters (3 & 5) are negative, and one cluster (4) is neutral.

Positive Clusters

The words that are used in the first two clusters are promising such as boost, jobs, and \$13 billion. These words give us an idea that the ailing economy that the UK is currently experiencing can get a boost of \$13 billion because of the Olympics as it will bring investment. This investment may be converted into the jobs and ultimately the GDP will

increase, as said by the Prime Minister David Cameron.

Negative Clusters

The words that are used in clusters 3 and 5 are discouraging. Words such as recession, empty seats, ghost town, bad etc. show us negative sentiments of the people about the holding of the Olympics in London. It is interesting to note that Greece and crisis are also in the same cluster. It seems like people believe London might experience similar financial turmoil as Greece experienced after it hosted the Olympic Games in 2004.

Neutral Clusters

The fourth cluster only has neutral words that are only related to Olympics in general such as Brazil, China, gold medals etc.

Cluster ID	Descriptive Terms	Frequency	Percentage
1	'ailing economy' +attend +business +cameron +class david +prime +bbc ailing +boost news +europe +minister leaders +osborne	...	145 10%
2	+uk +boost +effect foreign gdp +investment +govt 13bn happiness jobs +bring figures economy london2012 'economic boost'	...	241 16%
3	'economic growth' +'economic benefit' +britain +economic opening seats social +ceremony +'global economy' benefits +empty +recession +crisis +benefit +greece	...	371 25%
4	'olympic games' +'british economy' +brazil +british +country +money +olympic +spend +world countries gold greek medals +china +gold	...	407 27%
5	'economic impact' +ghost +impact +london people +town +city +video +bad +report cities +good hosting good +fuck	...	342 23%

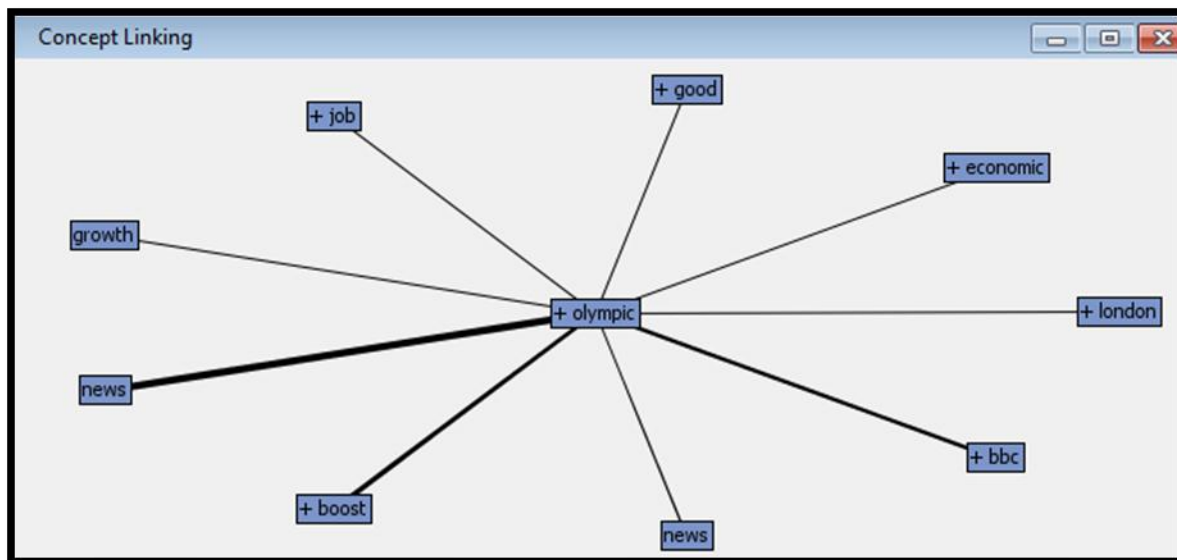
Display 3: Clusters formed

Concept Linking on Text Clusters

The concept linking diagram is created using same key word "Olympic" to give a comparison to the audience regarding the sentiments of the people.

Positive Clusters

The first diagram shows compilation of positive clusters. Here, words that are conceptually linked to the keyword are good, boost, job, growth. To enhance this result further, we can say that Olympic Games are associated with prosperity and job creation. These also bring boost and growth to the economy.

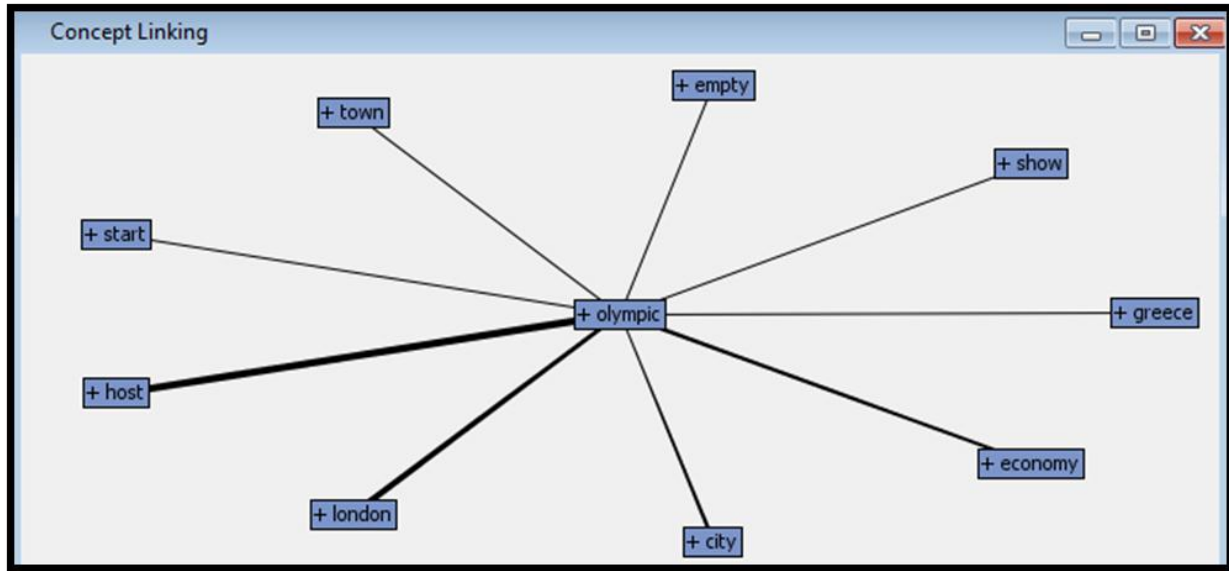


Display 4: Concept linking for positive text clusters

Negative Clusters

In this clusters, the keyword "Olympic" has been conceptually linked with more pessimistic words. Words such as London, empty and town suggest that people are referring to London as empty town because of Olympic Games event. Key word "Greece" is also used here that can be taken in a way that people are of the opinion that fragile economy of British will collapse just like Greek economy after Greece held Olympics in 2004 and went over budget by

60%.

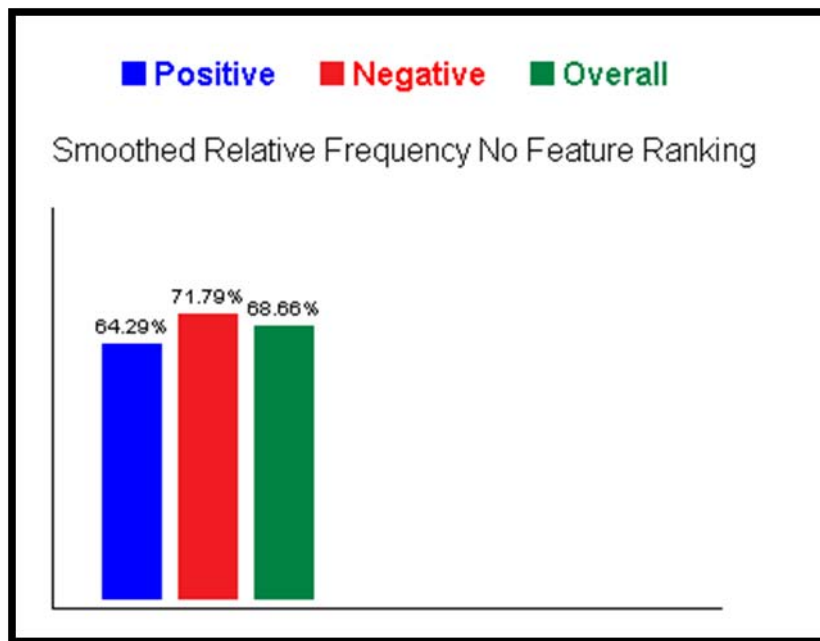


Display 5: Concept linking for negative text clusters

RESULTS USING SAS SENTIMENT ANALYSIS STUDIO

Sentiment Distribution

In order to corroborate our findings from SAS Enterprise Miner, we used SAS Sentiment Analysis Studio and built basic statistics model. Smoothed Relative Frequency No Feature Ranking model was selected as the best model. The validation result for this model showed that there were 71.8% negative sentiments and 64.3% positive sentiments found in the corpus with overall precision of 68.7%.

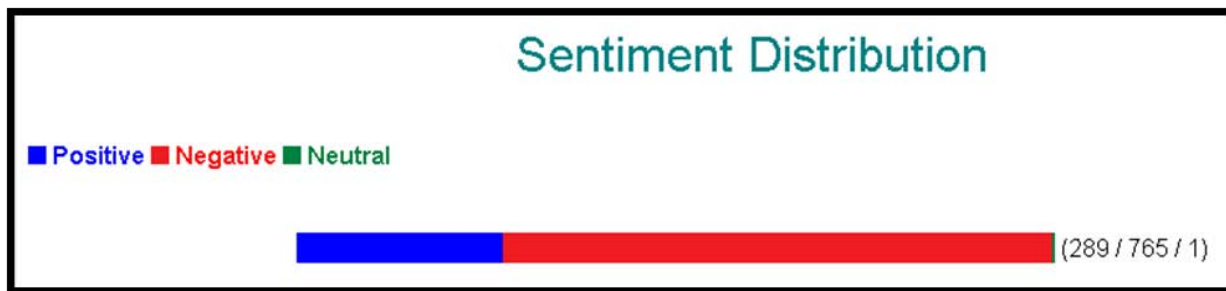


Display 7: The selected statistical model

The overall result in testing is strongly negative. Out of the 1,055 tweets that we tested, 765 (or 72.5%) were negative and only 289 (or 27.4%) were positive. This implies that most people think Olympic Games don't bring any financial benefit to the UK.

Text Result	Graphical Result
Results for selected folder:	
This directory is Negative	
Number of articles:1055	
Number of positive articles:289	
Number of negative articles:765	
Number of neutral articles:1	
Positive percent:27.39%.	

Display 8: Sentiment Distribution



Display 9: Sentiment Distribution in test data

Trend Analysis

We divided positive and negative tweets into three sections according to the weeks in which those were written. Week 1 represents the week before Olympics Games started. Week 2 includes the tweets of the week in which Olympics Games started. Week 3 has 10 days in it and represents the final week of Olympics.

	Sections
July 20 th – July 26 th	Week 1
July 27 th – August 2 nd	Week 2
August 3 rd – August 12 th	Week 3

Table 1: Division of tweets in weeks

The percent of tweets divided in positive and negative sentiments that we got in each week are:

	Positive	Negative
Week 1	35.66%	64.34%
Week 2	31.02%	68.98%
Week 3	36.03%	63.97%

Table 2: Division of positive and negative tweets in weeks

We created trend lines based on the results that we got and found some interesting facts. In week 1, before the Olympic games started, there was an overwhelming majority of negative tweets at 64.34%, whereas positive tweets were just 35.66%. This was the time when a majority of the people were furious with the huge government's spending on hosting the games. In week 2, when Olympics actually started, people became more negative as London became

an empty town and tourists started avoiding the city altogether. The instant improvement in the economy that the government claimed couldn't be seen anywhere that triggered people to become more pessimistic about the games. At this point, negative tweets soared to a new height of 68.98% whereas positive tweets fell to a new low of 31.02%. In week 3, when the Olympic games were concluding we saw a very interesting shift in the moods of the people. Negative tweets were at all-time low and positive tweets jumped to a new height. There were 36.03% positive tweets and 63.97% negative tweets. This is not a drastic change if we compare the overall result, yet it is interesting to see that people are getting less negative. There might be several reasons behind this shift such as the Paralympic games that were about to begin or new records set by Michael Phelps and Usain Bolt.

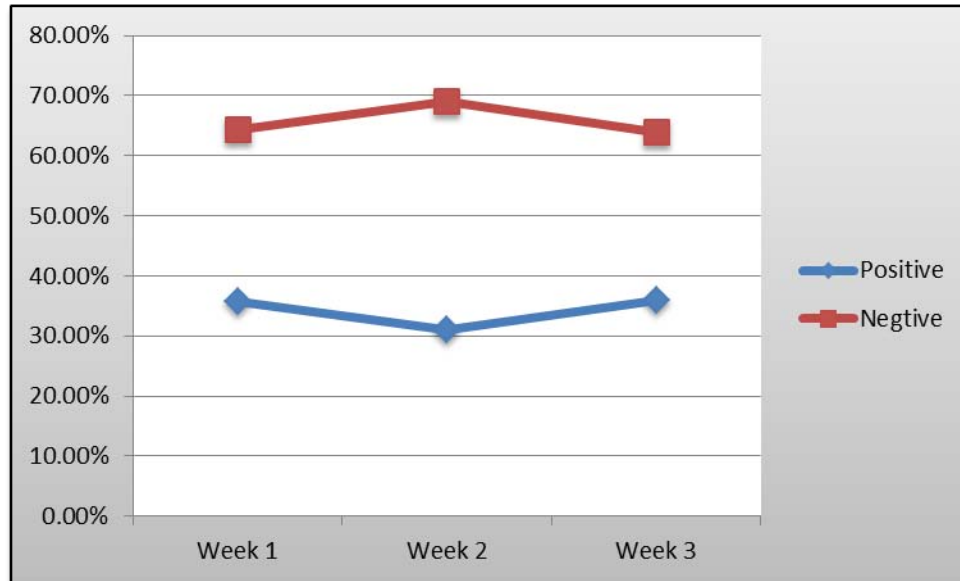


Figure 1: Sentiment Trend lines

CONCLUSION

The purpose of this research was to analyze the sentiments expressed by the people about the impact of the Olympic games on the economy of England. Based on the results we obtained from SAS Enterprise Miner, we found out that people have mixed feelings about the impact of the games on the economy. But, the results derived from SAS Sentiment Analysis Studio produced a clearer picture about the opinions of the people. It showed us that more people have reservations about the huge spending of the government on the Olympic games.

Social media is full of expressions these days, and we placed a limitation on us by just focusing on the data we collected from Twitter. With the knowledge of mankind expanding exponentially, it is difficult to get an accurate picture of the opinions around the world.

REFERENCES

- Kasimati E, Dawson P. Assessing the impact of the 2004 Olympic Games on the Greek economy: A small macroeconomic model. *Economic Modelling* 2009;26(1) 139-146.
- Larson D. All the easiest ways to search old tweets[Internet]. 2010[cited 2012 August 23]. Available from <http://blog.tweetsmarter.com/twitter-search/10-ways-and-20-features-for-searching-old-tweets/>
- Garla S, Chakraborty G. "%GetTweet: A New SAS® Macro to Fetch and Summarize Tweets," SAS Global Forum 2011; Paper 324-2011

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Name: Yu Fu
 Enterprise: Oklahoma State University
 Address: Department of MSIS, Oklahoma State University
 City, State ZIP: Stillwater, OK – 74075

Work Phone: 405-309-9356

E-mail: yu.fu@okstate.edu

Name: Shirmeen Virji

Enterprise: Oklahoma State University

Address: Department of MSIS, Oklahoma State University

City, State ZIP: Stillwater, OK – 74074

Work Phone: 901-857-1400

Fax: 901-471-4185

E-mail: shirmeen.virji@okstate.edu

Name: Dr. Goutam Chakraborty

Enterprise: Oklahoma State University

Address: Department of Marketing, Oklahoma State University

City, State ZIP: Stillwater, OK - 74074

Work Phone: (405)744-7644

E-mail: goutam.chakraborty@okstate.edu

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.