

Social Media Analytics: A Sentiment Analysis of a Hotel in New Delhi

Arvind Kumar¹ • Hansika Khurana² • Misha Matharu³

¹Head, Department of Commerce, LSR College for Women, University of Delhi

²Guest Faculty, LSR College of Commerce, University of Delhi

³Former Business Analyst, Verity knowledge solutions, Hyderabad

Email Id: dr.kumar.arvind@gmail.com¹ • hansika.khurana@gmail.com² • misha.matharu@gmail.com³

Abstract. Social Media Analytics analyses social media content and draw useful conclusions which helps to comprehend large amount of data and also give an entity competitive advantage. Through this study, we attempt to establish benefits of sentiment analysis to hospitality industry through case study of a single hotel; say Hotel ABC, New Delhi, using customer reviews on Trip Advisor. The performance of the hotel in various areas can be established through opinion that it creates. The following study brings into light areas where the hotel is doing exceedingly well and where it needs improvement. Such analysis can help the hotel to remain ahead in the game, take corrective actions and make customer experience satisfactory along with acquiring new customers.

Keywords: Social Media Analytics, Sentiment Analysis, Hotels, Hospitality, Mean, Median, Mode, Social Media, Online Reviews

1 Introduction

Social media is the social interaction among people in which they create, share or exchange information and ideas in virtual communities and networks. Andreas Kaplan and Michael Haenlein define social media as ‘*a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0 and that allow the creation and exchange of user-generated content.*’

Furthermore, social media depends on mobile and web-based technologies to create highly interactive platforms through which individuals and communities share, co-create, discuss, and modify user-generated content. They introduce substantial and pervasive changes to communication between organizations, communities, and individuals.

Social media offers advantages over traditional media on parameters including quality, reach, frequency, usability, immediacy, and permanence. According to recent research, internet users continue to spend more time with social media sites than any other type of site. There are several dozens of social networks and their numbers keep growing by the day. The most popular ones are Facebook, Google, LinkedIn, Skype, Twitter and Trip Advisor.

Classification of social media

Social media technologies take on many different forms including magazines, Internet forums, weblogs, social blogs, micro blogging, wikis, social networks, podcasts, photographs or pictures, video, rating and social bookmarking.

A few types of Social Media are as follows:

- i. Collaborative Projects (for example, Wikipedia)
- ii. Blogs and Micro Blogs (for example, Twitter and Tumblr)
- iii. Content Communities (for example, YouTube and Daily Motion)
- iv. Social Networking Sites (for example, Facebook)
- v. Virtual Game-Worlds (e.g., World of Warcraft)
- vi. Virtual Social Worlds (e.g. Second Life)

Internet usage effects

An increasing number of scholars have sought to study and measure the impact of social media. A 2010 study by the University of Maryland suggested that social media services may be addictive, and that using social media services may lead to a "fear of missing out," also known as the phrase "FOMO" by many students. It has been observed that Facebook is now the primary method for communication by college students in the U.S. According to Nielsen, global consumers spend more than six hours on social networking sites. "Social Media Revolution" produced by Socialnomics author Erik Qualman contains numerous statistics on social media including the fact that 93% of businesses use it for marketing and that if Facebook were a country it would be the third largest. Several colleges and universities such as Harvard, Johns Hopkins, Columbia and Stanford among others have even introduced classes on best social media practices, preparing students for potential careers as digital strategists.

Effects of using social media for news purposes

Just as television turned a nation of people who listened to media content into watchers of media content, the emergence of social media has created a nation of media content creators. According to 2011 Pew Research data, nearly 80% of American adults are online and nearly 60% of them use social networking sites. More Americans get their news via the Internet than from newspapers or radio, as well as three-fourths who say they get news from e-mail or social media sites updates, according to a report published by CNN. The survey suggests that Facebook and Twitter make news a more participatory experience than before as people share news articles and comment on other people's posts. According to CNN, in 2010 75% of people got their news forwarded through e-mail or social media posts, while 37% of people shared a news item via Facebook or Twitter.

1.1 Social Media Analytics

It refers to Measuring, Analyzing and Interpreting interactions and associations between people, topics and ideas. Social media analytics is a powerful tool for uncovering customer sentiment dispersed across countless online sources. This analysis is often called Social Media Listening or Online Listening. *The analytics allow marketers to identify sentiment and identify trends in order to better meet their customers' needs.* There have been significant examples where companies, such as Whirlpool, Royal Bank of Canada and JetBlue have used such analytics tools to engage customers in response to their feedback.

Social media analytics gathers information from social networking streams such as Facebook and Twitter, as well as blogs, news feeds, and comments from rich media sites, to better understand customer attitudes and characteristics. Reviews of products and services online are often considered by customers who would be potential buyers of those products or services. For example, reviews of restaurants on Zomato or product reviews on Flipkart.

It creates new opportunities for companies that want to engage better with customers. Real engagement, which binds customers to companies, can drive revenue gains and reduce the costs associated with customer churn.

Considering all the benefits that social media can provide it has all the attributes that will be helpful in marketing and management. This is particularly important for *hospitality and tourism area field* since they are people professions. Social media provides opportunities to manage and present business content that is most appealing to customers. It helps in identifying

core values that draw clients and customers to business. Most companies would benefit from a cohesive strategic plan that will integrate social media efforts with overall marketing effort.

Engaged customers spend more money and return back to business more often. It also helps in brand awareness and building brand equity. Sentiment Analysis as a part of social media analytics is very important to today's businesses. It is essential to understand and analyze opinion and take action regarding the content generated by the customers. It helps to make better business strategies, product development, improving customer satisfaction and brings into limelight the existing loopholes. Therefore, to understand the sentiment that the hotel industry generates among its customers, we have used Trip Advisor, a travel website providing travel-related reviews as well as interactive travel forums.

TripAdvisor is the world's largest travel site enabling travelers to plan a perfect trip. In addition to hotel bookings and reservations, it also has more than 170 million reviews and opinions covering more than 4 million accommodations, restaurants and attractions. This helps the customers make informed decisions regarding their travel and stay plans. Most visitors to the website have indicated that they have found the reviews posted online as *very helpful*.

Managing hotel reviews, both positive and negative is certainly a priority for managers. Sites like TripAdvisor are constantly on the radar of managers to assess customer experience. Thus, Sentiment Analysis through TripAdvisor can help a Hotel manager improve the areas which are generating negative sentiments among customers and propagate the strengths. In this paper, through case study method we have conducted sentiment analysis for a single hotel, Hotel ABC, New Delhi through customer reviews on TripAdvisor for a period of 3 months - May, June and July.

2. Literature Review

In a review by Martin Soler and Josiah Mackenzie (2011), it was stated that online reviews received a lot of attention in the hotel industry from a reputation management perspective; these helped in setting the direction of marketing and advertising campaigns. It was found that the attention devoted to online reputation management for hotels has been remarkable. Hotel consultants and agencies were looking to educate hoteliers to encourage guests to share their experience online, and to draft proper management responses to a review.

The E-Tailing group, in its study, found that 89% of people admitted to have been influenced by reviews while making a purchase decision. Also, in an Econsultancy survey, a majority of online shoppers said they trusted recommendations from people they knew, while most of them trusted opinions of unknown users.

A careful study of a hotel's online reviews revealed exactly what guests liked and didn't like. With this information, it could be known what to avoid and what promote.

Walter Kasper and Mihaela Vela (2011) found that travel planning and booking on the web had become one of its most important commercial uses. With the rise of the Web-2.0 user-generated reviews, comments and reports about their travel experiences played an increasing role as information sources. This was true especially for hotel bookings. User reviews were found to be more trustworthy than information printed in hotel catalogs. User reviews tended to be unbiased and reflected actual experiences of the customers.

In a study on Hospitality Industry by Attensity (2011), it was found that travelers were becoming increasingly vocal and were willing to share their experiences with hotels and travel providers. All ratings on various attributes and sub-attributes of a hotel provide valuable insights to hoteliers to enable them to improve guest experience and differentiate them from their competitors. The report stated that customers were more likely to share their experiences on social media forums like Facebook, Twitter or TripAdvisor.

3. **Research Methodology**

In order to analyze Social media content generated for a Hotel, we have used case study approach. The approach is useful for an in-depth analysis for a limited number of events; in this case reviews by customers about a hotel on specific parameters. The approach used in a single case study can then be further extended to the other hotels and inter hotel analysis.

A list of 17 Attributes has been established through a thorough reading of the reviews on TripAdvisor. Research articles have been used to make the list comprehensive. Whereas a multitude channels are available to conduct sentiment analysis for a hotel, we have limited the research to TripAdvisor as it is considered one of the most popular travel forums. Reviews for a period of 3 months - May, June and July 2014 for Hotel ABC have been ranked on 17 attributes mentioned above using the following ranking system:

Ranking System

2 - Extremely good

1 - Good

0 - Neutral comment or no mention

-1 - Below expectation or disappointing

-2- Very disappointing or strong emotion expressed against the hotel regarding a particular attribute of the hotel

Statistical Techniques

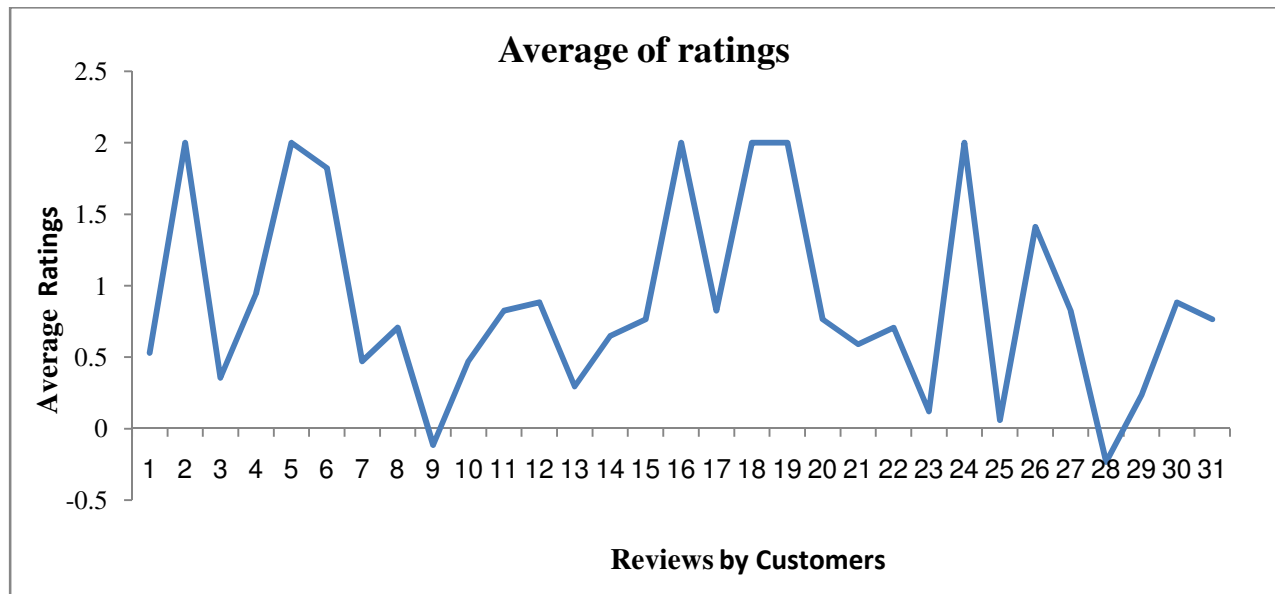
- i. Average or mean - "The statistical mean refers to the mean or average that is used to derive the central tendency of the data in question. It is determined by adding all the data points in a population and then dividing the total by the number of points. The resulting number is known as the mean or the average."- Techopedia
- ii. Mode - The mode of a distribution with a discrete random variable is the value of the term that occurs the most often. Standard Deviation - The standard deviation is a numerical value used to indicate how widely individuals in a group vary. If individual observations vary greatly from the group mean, the standard deviation is big; and vice versa.

Simple statistical techniques have been used to analyze customer reviews for Hotel to give an insight into the strengths and weaknesses of the hotel.

4. Findings of the Study

The reviews given by customers are not only important for brand reputation but also for new customer acquisition. A Study conducted by Phocuswright on behalf of TripAdvisor in 2012 concluded that 53% of respondents will not book a hotel that doesn't have reviews on the site and 87% agree that TripAdvisor hotel reviews "*help me feel more confident in my decisions.*"

The mean of each review over the various attributes can be interpreted as an overall sentiment for the customer. The graph below indicates an average rating of each customer review in the stated 3 months. It indicates that not many customers have given an overall bad review to the hotel with only few dips below 0, which means disappointing performance. Also, many spikes indicate that there are many reviewers who have given an overall positive feedback on various attributes.



Source: The Researcher

Standard Deviation

Standard Deviation is a measure of variability used in statistics. A low standard deviation indicates that the data points tend to be very close to the mean, whereas high standard deviation indicates that the data is spread out over a large range of values. The values of each review are difficult to interpret but range is between 0.7 - 1.5. Certain reviews with 0 standard deviation shows that throughout the review the sentiment has been consistent.

Co-efficient of Variation

Co-efficient of variation is used to summarize and better analyze data. It is defined as the ratio of the standard deviation to mean. When the value of coefficient of variation is higher, it means that the data has high variability and less stability. When the value of coefficient of variation is lower, it means the data has less variability and high stability.

Mean of all the means of reviews = 0.888

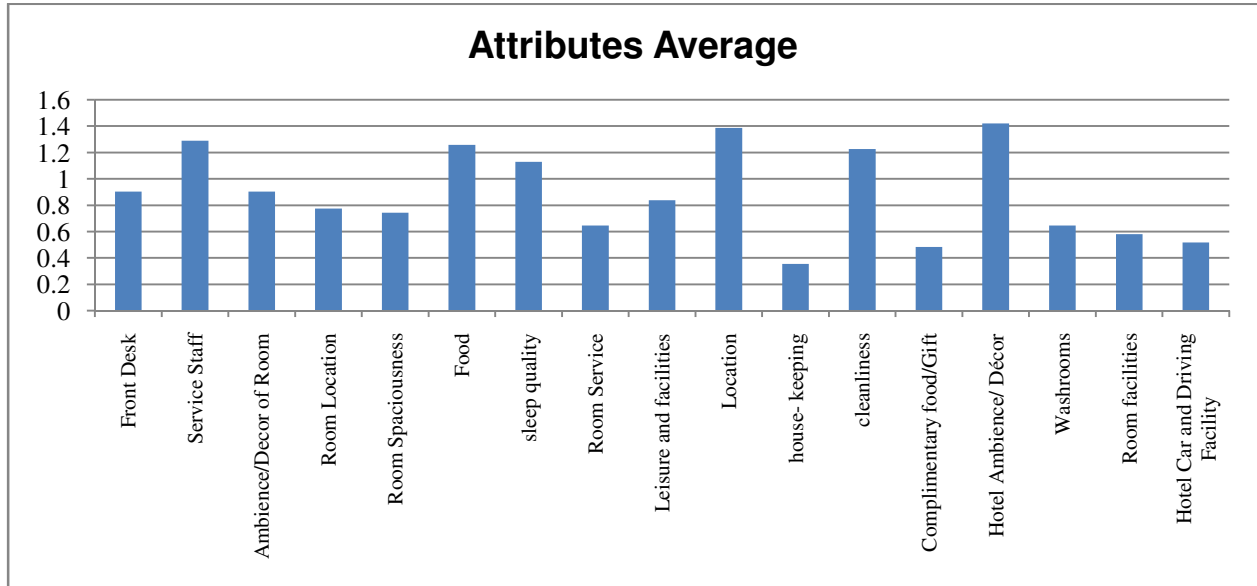
Standard Deviation of reviews = 0.682

Co-efficient of Variation= 77%

Thus, in this case stand alone mean that is 0.88 cannot be considered representative as co-efficient of variation is high.

Averages of Attributes

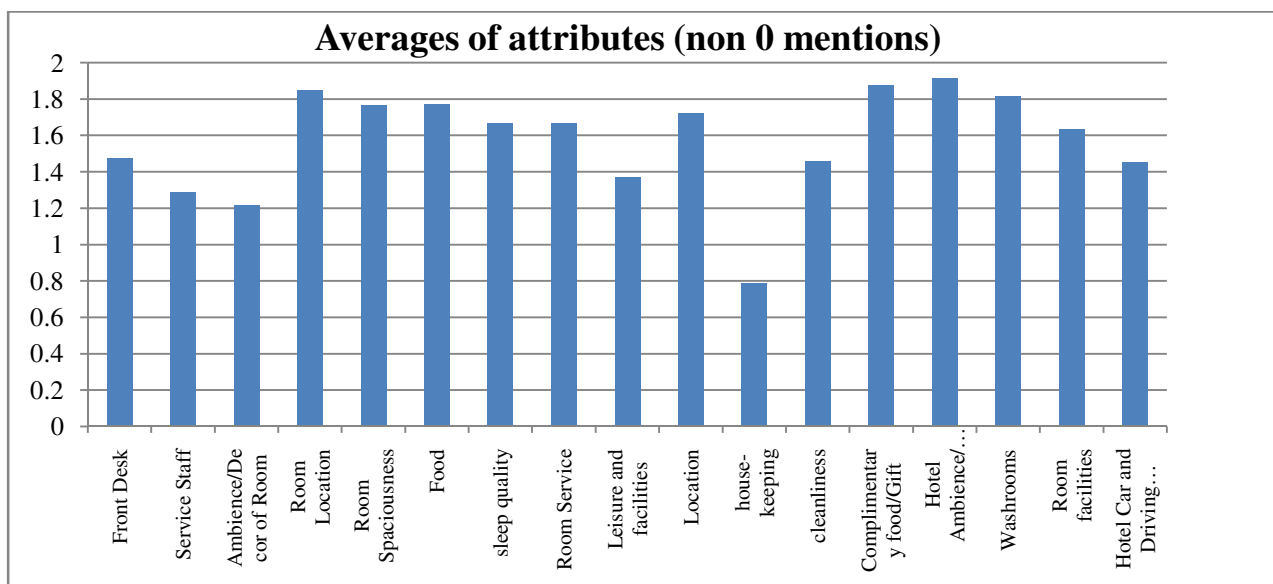
The ratings for each attribute by all reviewers over the period of 3 months gives an insight into general sentiment about each attribute.



Source: The Researcher

From the graph it can be observed that the highest ratings have been for Hotel Decor and Location approximately around 1.4 suggesting a positive sentiment. The lowest rating has been for House-keeping approximately 0.4 implying a neutral to slightly positive sentiment.

A still better representation would be to consider non 0 mentions while averaging.



If we consider only non 0 mentions, the averages of attributes have improved drastically. The highest average rating is still of Hotel Ambience/ decor but it has gone upto 1.9 which implies exceeding expectation. The lowest rating is of House-keeping that is 0.8, which still means positive sentiment. This graph also shows that the hotel is performing well on all attributes with general positive sentiment as shown by the graph

The mode is the value that appears most often in a set of data. In this case, we have calculated the mode for each attribute.

Attributes	Mode
Front Desk	2
Service Staff	2
Ambience/Decor of Room	2
Room Location	0
Room Spaciousness	0
Food	2
sleep quality	2
Room Service	0
Leisure and facilities	2
Location	2
house- keeping	0
Cleanliness	2
Complimentary food/Gift	0
Hotel Ambience/ Décor	2
Washrooms	0
Room facilities	0
Hotel Car and Driving Facility	0

Thus, 2 has been the most frequent rating which shows that that hotel has performed well.

The attributes with the maximum number of non 0 mentions shows that these attributes are important to customers and are frequently mentioned by them in their reviews.

Service Staff	31
Cleanliness	26
Location	25
Ambience / Decor of Room	23
Hotel Ambience/ Décor	23

Another finding worth noting is that Service staff also has received the maximum -2 ratings which suggest much lower customer expectations.

5. Conclusion and Recommendations

Social media can have both positive and negative impact on not only smaller businesses but also on franchise corporations. Actions become necessary and timely response to the reviews when they are genuine can result in a big payoff. This becomes important when brand image is concerned. Many hospitality businesses have utilized social media beneficially. Whether the reviews were positive or negative, they have adequately responded to their advantage. With the rapidly changing technology it is impossible for any hospitality operator to ignore the social media.

After analyzing the attributes of Hotel ABC, we can conclude that the reviews of the hotel have mostly been positive. The overall sentiment of the hotel is consistent showing more than satisfactory performance. However, we cannot consider the overall mean as an absolute indicator of a positive sentiment. This is due to the fact that there is a high Coefficient of Variation at 77% indicating high variability in responses.

Whether we consider the 0s or not, Hotel Décor and Location remains the most highly rated attribute. The hotel has very close proximity to some of the best malls and shopping complexes in Delhi making the location ideal. However, house-keeping has received the lowest ratings which range from 0.4 – 0.8 but yet shows a positive sentiment. Therefore, the hotel has been successful in keeping its customers satisfied showing a positive sentiment. Hotel Décor and Location is a major strength of the hotel while its House-Keeping is a weakness which needs to be worked upon. 9 out of 17 attributes have received a rating of 2, which is the highest rating that an attribute can get. This shows that the hotel is meeting the expectations of its customers. Certain attributes have been pointed out as the most essential while measuring customer sentiment namely Service Staff, Cleanliness, Location, Ambience of Room and Ambience of the Hotel.

One of the most important attributes which has been mentioned in majority of the reviews is Service Staff. This is the attribute which has received the highest number of -2 ratings showing customer disappointment. This is a major concern area for the hotel and needs to be strengthened. Although the hotel replies to each review, whether positive or negative, it is yet to be ascertained whether this attribute has been looked into or not. Although, we have tried to present a comprehensive analysis on Hotel ABC, our study has certain limitations

- i. We have presented a case study of a single hotel. By including more hotels, we could have provided a detailed inter-hotel analysis and comparison.
- ii. We have used a single source i.e. TripAdvisor for our secondary information needs. There are other platforms like Yelp.com which also provide hotel reviews and ratings.
- iii. The ratings could be biased since the interpretation of reviews could be subjective depending from person to person.
- iv. We have taken reviews for a period of 3 months. Owing to the limited number of reviews, we have not been able to present a frequency distribution.
- v. The number of 0s does not always mean a fact stated or not mentioned. It could also mean a sum of the +1s and -1s while calculating the values pertaining to various sub-attributes. For example, the attribute Room could have a rating of 0 due to the fact that one of its sub-attributes, say Room Décor has a rating of +1 while another sub-attribute; say Room Location has a rating of -1.

6. **Social Media Analytics Pitfalls**

Many marketing professionals seem to feel the need to embrace “big data” solutions and social media, and to become more data driven in their decision making. The danger in doing so is that they will become overly reliant on the data from these systems and on marketing through social media channels. While social media is a viable source of information and valid marketing channel, social media marketing and analytics tools also have several drawbacks and caveats. These include problems with bias, message drift, and analyzing non-textual content. Marketing professionals should consider social media data and channels as some of many tools in their marketing toolkits but not the only ones. Traditional market research and marketing channels cannot be ignored in favor of social media.

Age and Geographic Bias: At present, an acknowledged bias in social media data can affect social media analytics’ output. Social media use is more likely amongst younger people and in certain geographies. There are geographic biases in social media as well. For example, more activity in social media exists on the East and West Coasts of the United States than in other places, causing this data to have an outsized effect. Some social media analytics vendors normalize their data in an effort to reduce these biases. Even so, marketing professionals need to consider bias when making decisions. When marketing to more general groups, social media

data can only be viewed in light of data from other sources that does not have these biases. It is worth noting that it is possible that this bias will disappear over time. That is, however, likely to be far into the future, and marketing professionals need to approach social media data with these biases in mind.

A Downside to Amplification: One of the reasons that social media is such a cost-effective marketing tool is amplification. Using network effects, social media can take an investment in content and messaging and spread it with little to no incremental costs. *Amplification has a serious drawback though: message drift.* Social media followers often simply repost—in other words, share—posts and content. That’s preferred since it amplifies without changing the message. Sometimes people will rephrase or shorten a message, inadvertently changing the meaning. When viewers comment on a post, blog, or video, they can change the basic message encoded in the post, even interjecting entirely new messages. With social media, it is difficult to control the message without constant attention.

Non-textual Information Presents a Problem: Another potential problem with social analytics platforms is their reliance on text data. Most sentiment analysis and those which demand signal tools, for example, are unable to directly analyze the growing amount of image, video, and audio content. While many of these products can provide useful information derived from the metadata and comments associated with social content, they cannot decode the actual meaning of the non-textual content itself where the majority of the information is contained. This means that the growing body of content contained in video blogs, on video sharing sites such as YouTube, on photo sharing sites such as Instagram and Pinterest, or in podcasts as well as much of the content on Facebook do not yet figure prominently into social media analysis.

References

Broder, A., Fontoura, M., Josifovski, V., Riedel, L.(2007).: Semantic approach to contextual advertising.In: Proc. of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Amsterdam, The Netherlands Online advertising: behavioural targeting available at

http://www.networkadvertising.org/pdfs/NAI_Beales_Release.pdf

<http://www.oracle.com/us/corporate/analystreports/es-social-landscape-1936813.pdf>

<http://yourstory.com/2014/03/social-media-analytics-market/>

http://www.hicss.hawaii.edu/hicss_47/bestpapers/Social%20Media%20Analytics%20and%20Business%20Value_c_%20A%20Theoretical%20Framework%20and%20Case%20Study.pdf

<http://www.pwc.co.uk/hospitality-leisure/publications/the-social-media-opportunity-for-hotels.jhtml>

<http://ro.uow.edu.au/cgi/viewcontent.cgi?article=1281&context=commpapers>

<http://scitechnol.com/2324-9129/2324-9129-1-e102.pdf>

http://www.tripadvisor.co.uk/PressCenter-c4-Fact_Sheet.html

<http://www.techopedia.com/definition/26136/statistical-mean>

<http://searchdatacenter.techtarget.com/definition/statistical-mean-median-mode-and-range>

http://stattrek.com/statistics/dictionary.aspx?definition=standard_deviation

<http://en.wikipedia.org/wiki/TripAdvisor>

<http://www.wihphotels.com/publication/guide-to-using-social-media-and-review-analytics-for-hotel-marketing/>

<http://www.proceedings2011.cla-conf.info/pliks/25.pdf>

http://www.attensity.com/assets/industry_report_hospitality.pdf