Social Media Analysis: A Business Imperative

Chantelle Stephen
H00200692

August 2015

Computer Science
School of Mathematical and Computer Sciences

Dissertation submitted as part of the requirements for the award of the degree of MSc in Computer Systems Management
Declaration

I, Chantelle Stephen confirm that this work submitted for assessment is my own and is expressed in my own words. Any uses made within it of the words of other authors in any form (e.g., ideas, equations, figures, text, tables, programs etc.) are properly acknowledged. A list of references employed is included.

Signed: ......................................................................................

Date: ...........................................................................................
Abstract

In the recent years, social networking for personal and business purposes is a common phenomenon. Leveraging and manipulation of data or information obtained from social platforms, has been a trend that has been gaining increased popularity. Corporate leaders have realized the power of social media with respect to gaining insights about current market situations and trends, communicating with consumers and business partners, increasing profits as well as to take strategic actionable decisions. However, certain organizations are still hesitant to venture into this comparatively new marketing technique due to a number of concerns and fears related to the topic. This research is being conducted within the United Arab Emirates, where technological advancements are in match to international standards and is home to a rich diversity of culture speaking various languages. The study is focused on understanding five major topics which includes; the influence of social media within the region, various organizations that adopt and do not adopt social media analysis, concerns and limitations related to social media marketing and its analysis, significance of multi-lingual analysis as well as how various sectors of industries could benefit from social network analysis. This paper also proposes effective steps organizations could adopt, in order to efficiently manipulate and leverage social networks by strategically overcoming concerns and limitations related to it as well as increasing consumer retention and satisfaction. Issues related to the limited awareness of social media analysis and its benefits among the people, causes a tendency to overlook the gains that could be obtained from this source; these issues are also researched and elaborated in this study.

Keywords: Social media, Social media analysis, Multi-lingual analysis, Social media marketing.
Acknowledgement

To begin with, I thank *The Almighty* for giving me patience, courage, time and endurance to pursue and complete this research. My supervisor and mentor, Dr. Hind Zantout, without whose constant guidance, review and support this study would not have been possible. Her innovative ideas complimented my research topic and helped improve the quality and context of my work. My parents, who have constantly supported me, both physically and emotionally in order to complete this study. My peers, who have enlightened me with their ideas, reviews and feedback on the research as well as the report regarding readability, clarity and understandability which helped me improve the report as a whole. Last but not the least, I would like to thank all my professors, colleagues and classmates whose ideas influenced one topic of the research or the other and was continuous process throughout this post-graduate course.
# Table of Contents

Declaration ........................................................................................................................... I
Abstract ............................................................................................................................... II
Acknowledgement ............................................................................................................... III
List of Figures and Tables ..................................................................................................... VI

Chapter 1 – Introduction ........................................................................................................ 1
  1.1 Aims and Objectives ..................................................................................................... 2
  1.2 Structure of the report ................................................................................................. 3

Chapter 2 – Literature Review ............................................................................................... 5
  2.1 Social Media Analytics (SMA) ..................................................................................... 5
    2.1.1 Terminologies related to SMA ............................................................................. 5
    2.1.2 Top Grossing Social Media Platforms and their leveraging .................................. 6
  2.2 Analysis and visualization tools that can be used for SMA ............................................ 8
    2.2.1 Visualizing brand activity and personality ......................................................... 8
    2.2.2 Sentiments and opinion visualization ................................................................... 9
    2.2.3 Common other visual representations .................................................................. 11
  2.3 Benefits of Social Media Analysis ............................................................................... 12
    2.3.1 Customer related benefits ................................................................................... 12
    2.3.2 Business related benefits .................................................................................... 13
    2.3.3 Organizational related benefits .......................................................................... 14
  2.4 Concerns related to employ Social Media Platforms ..................................................... 15
    2.4.1 User related concerns ........................................................................................ 15
    2.4.2 Business related concerns .................................................................................. 16
  2.5 Limitations of Social Media Analysis .......................................................................... 18
  2.6 Multi-lingual analysis – A critical requirement within Social Media Analytics .............. 21
    2.6.1 Relevant research and studies related to multilingual analysis ......................... 22
  2.7 Effective utilization of Social Media Analytics for organizations .................................. 23
    2.7.1 Background Research ......................................................................................... 23
    2.7.2 Aims and Objectives ........................................................................................... 24
    2.7.3 Building on a strategy ........................................................................................ 25
    2.7.4 Evaluate and measure ........................................................................................ 26
    2.7.5 Take Actions ...................................................................................................... 27
  2.8 Social Media Analytics in Action – Success Stories ...................................................... 27
Appendix

Chapter 3 – Research methodology ................................................................................................................. 29
  3.1 Research problem .............................................................................................................................................. 29
  3.2 Research questions .............................................................................................................................................. 29
  3.3 Methodology ....................................................................................................................................................... 30
    3.3.1 Stage 1: Quantitative Research using surveys and polls ................................................................................ 31
    3.3.2 Stage 2: Qualitative Research using Focus Group Discussions ................................................................. 32
  3.4 Research Analysis Plan ....................................................................................................................................... 33

Chapter 4 – Research Results .................................................................................................................................. 34
  4.1 Quantitative data analysis – Survey and polls ............................................................................................... 34
    4.1.1 Phase 1: Influence of Social media within the U.A.E .............................................................................. 34
    4.1.2 Phase 2: Social media analytics and leveraging in organizations within the U.A.E ............................ 37
    4.1.3 Phase 3: Participant’s understanding of social media analysis and its importance ............................ 41
    4.1.4 Hypothesis Testing ....................................................................................................................................... 44
    4.1.5 Poll results .................................................................................................................................................... 47
  4.2 Qualitative data analysis – Focus Group Discussions ..................................................................................... 49
    4.2.1 Day 1: Focus Group Discussion regarding SMA and its uses and concerns in U.A.E ....................... 49
    4.2.2 Day 2: Focus Group Discussion regarding SMA limitations and visualization qualities .................. 52
  4.3 Framework for major Industry-Sector’s and their benefits using SMA .......................................................... 54

Chapter 5 – Evaluation ............................................................................................................................................ 61
  5.1 Limitations and Future Work ......................................................................................................................... 62

Chapter 6 – Conclusion ........................................................................................................................................... 64

References ................................................................................................................................................................. 66

Bibliography ............................................................................................................................................................ 72

Appendix .................................................................................................................................................................... 73

Appendix A – Survey Form ....................................................................................................................................... 73
Appendix B – Polls ..................................................................................................................................................... 78
Appendix C – IBM SPSS (Hypothesis Testing) ..................................................................................................... 79
Appendix D – Focus Group Discussion .................................................................................................................. 82
Appendix E – R Programming .................................................................................................................................. 83
Appendix F - Images ................................................................................................................................................... 90
List of Figures and Tables

Figure 1- Brandwatch dashboard showing different social media activities of a particular brand .......... 8
Figure 2- Heat Map created using Twitter Maps .................................................................................. 9
Figure 3- Sentiment Visualization using Sentiment Viz for UAE .................................................................. 9
Figure 4- Sentiments depicted over a timeline using Sentiment Viz for Etisalat ................................. 10
Figure 5- Sentiment word cloud over changing oil prices collected over 2014-15 by US Government ... 10
Figure 6- Cluster-node visualization created using online tool SociLab ............................................. 11
Figure 7- Bubble Chart example created using IBM ManyEyes for popular digital camera models .................................................................................. 11
Figure 8- New Look Middle East Instagram Post .................................................................................. 21
Figure 9- Enlarged view Arabic comments on New Look post shown in the image above ................. 22
Figure 10- Building blocks for effective utilization of social media networking and its analysis .................. 23
Figure 11- Mixed Research Analysis method proposed for the study ....................................................... 30
Figure 12- Pie-Donut chart depicting the social media engagement within the U.A.E ......................... 34
Figure 13- Common types of posts seen on social networks among the participants .......................... 35
Figure 14- Frequencies of languages other than English commonly seen within the U.A.E .............. 35
Figure 15- Pie chart showing the purchase-frequency on social networks among participants ............... 36
Figure 16- Most influential mediums that convince consumers to invest in services or products ........ 36
Figure 17- Percentage of large scaled companies using SMA ................................................................. 37
Figure 18- Percentage of medium scaled companies using SMA ............................................................ 38
Figure 19- Percentage of small scaled companies using SMA ............................................................... 38
Figure 20- Pie Chart showing various industry sectors and their engagement in SMA ......................... 38
Figure 21- Participant’s response on the methods their companies use for SMA ................................. 39
Figure 22- SMA effectiveness in organizations ...................................................................................... 39
Figure 23- Various reason(s) provided by participants for their organizations not using SMA .............. 40
Figure 24- Percentage of participants with knowledge on data analysis on social data ....................... 41
Figure 25- Awareness of Social media analysis, benefits and tools among the participants ....................... 42
Figure 26- Response shown by participants with knowledge about SMA .............................................. 42
Figure 27- Response shown by participants with less knowledge about SMA .......................................... 43
Figure 28- Response shown by participants with no knowledge about SMA .......................................... 43
Figure 29- IBM SPSS Hypothesis Test Results showing count, mean and standard deviation ........... 45
Figure 30- IBM SPSS Hypothesis Test Results showing 2 tailed p value and mean difference .......................... 45
Figure 31- Participant’s views on various purposes of social media networks ........................................ 47
Figure 32- Participant’s view on department responsible for Social media engagement ..................... 48
Figure 33- Word Cloud creating using R Programming showing words associated with #Expo2020 .......... 52
Figure 34- Word cloud created using R Programming showing sentiments related to #Petrol .................. 53

Table 1- Top 5 social networks and types of information that can be obtained from them .................. 7
Table 2- Companies classified according to their number of employees .............................................. 34
Table 3- Focus Group Discussion participants and their related designations and sectors .................. 49
Table 4- UK Classification of sectors and related SMA benefits ........................................................ 60
Chapter 1 – Introduction

The advancement of the Web 2.0 and the popularity of smart phones has helped form a technological platform for sharing knowledge, effective communication as well as content generation; thereby forming the social media. The amount of data accumulated on the social media increases day by day and the data produced becomes bulkier every year. Social media is about the interaction between people and the information that they share on social networks such as Facebook, Instagram, Twitter and similar platforms. With big data being the center of attraction between researchers and academia in the technology sector, social media analysis (SMA) falls under this category and is a trending field. SMA can be referred to the study of human interactions, opinions and sentiments on various social media platforms, forums, blogs and related online facilities. The most basic or general fundamentals of SMA is capturing, understanding and presenting data.

With the major acceptance of social media in the last decade and being one of the major modes of communication and information provider, organizations are forced to think of business models to keep up and benefit from this trend. Many researches and studies in the field show an integration of business analytics along with social media analysis will help organizations gather information in order to improve customer experience and might even enable a 360 degree view from a customer’s perspective thereby gaining insights on customer requirements and changing trends, thereby improving sales and gaining more profits. Majority of the leading organizations from various sectors use social media for advertising and marketing purposes, and similarly gather essential information from such platforms to enable actionable decision making. There are major concerns related to the Return of Investment (ROI) within this form of analytics. Constant worry is caused about the financial returns social networks could provide with comparison to the investment in this form of marketing or business. What organization’s need to realize is that just like a business phone, which is used to connect to clients or potential clients in order to increase sales; social networking can be used for the same concept in a faster and more effective method. Information on what consumers think about the brand, opinions on the products and services as well as information gathering about competitors can also be obtained using these platforms. Information on improving customer experience, retaining customers, competitor’s engagement and the various changing trends in the business world, are the major returns that SMA can provide in order to increase sales and improve profits. Sentiment Analysis is understanding public’s opinions and sentiments over a various topic. For example in a blog various comments are posted, sentiment analysis can determine positive and negative feedback on that particular topic. Similarly leveraging social media networks by performing various analytics can provide valuable insights on various topics which any form of organization can benefit from.
Though English has been accepted as the universal language for communication, it is not the native language for majority of the world’s population. With social networks gathering popularity all over world, number of languages other than English can be seen prominently on these networks which may vary from one location to another. For example, within the Middle East where English is not the native language and with a large number of expatriates in the region; people tend to post on social networks especially in Arabic and various other languages. Majority of the social networks support different language and the need for multilingual analysis must be realized in order tend to all types of customers where language must not pose as a limitation. Organizations in a geographical position like within the United Arab Emirates (U.A.E) have clients of different cultures and using different languages. Sentiments expressed in other languages cannot be ignored and must be analyzed in order obtain high quality results and better insights.

1.1 Aims and Objectives

The aim of this research is to understand about the influence of social media, the depth of its analytics as well the importance of SMA in a country like the United Arab Emirates (U.A.E). Keeping in mind its geographical position, technological advancements, large number of expatriates and most importantly being host of the Expo 2020. This case study being conducted in the U.A.E will not only show the importance of social media marketing and analytics in this region, but also will highlight concerns and limitations related to the topic as well as how various sectors of the industries can benefit from SMA.

The major objectives of this study are:

I. Gather relevant information by studying previous research and work related to the topic in order to gain more understanding within this context.

II. Perform a quantitative research using participants within the U.A.E in order to gain deeper understanding with respect to the following topics. Detailed information about the research methods and target audience can be seen in section 3.3.1 of the report.

- The influence of social media within the region.
- The need of multilingual analysis within the region.
- Consumer’s purchasing-activities on social networks.
- Organizations based on the sizes, using SMA as part of the business activities
  - If leveraging social networks, methods used for SMA and how effective is it for the organizations?
  - If not utilizing social media marketing and its analytics, reasons related to it?
- Understanding the awareness of social media analytics and its benefits within the region.
• Realizing participant’s understanding about the importance of SMA within an organization.
• Conduct a hypothesis test from the data collected in order to find if there is a relationship between a participant’s knowledge on SMA and his/her perception about how important SMA is for an organization.

III. Conduct a qualitative research using participants who are professionals and have knowledge with respect to the research topic. This study will help obtain better knowledge with respect to the following questions as well as to support the quantitative research.

• What are the general concerns related to SMA within the U.A.E?
• How to eradicate the stereotype “SMA is only for the Marketing team”?
• What is the need for multi-lingual analysis within the region?
• How can the upcoming event “Dubai Expo 2020” benefit from SMA practices?
• What are the limitations of social media analytics and its visualization?

IV. Build a framework for major industry sectors and how they can benefit from SMA. Justify the points with the help of related studies in those sectors as well as with the use of local examples.

1.2 Structure of the report

The research report is divided into 6 chapters and is followed by a reference list, bibliography and appendix. The chapters have been subdivided for the ease of user readability and understanding. The chapter are as follows:

Chapter 1: Introduction – A brief introduction about the research topic. The aim and objectives of the study as well as the structure of the report has been explained in this chapter.

Chapter 2: Literature Review – A critical review on related research and studies within this topic have been made in order to show the relevance of the topic in today’s world. The chapter discusses in depth about the topics of SMA, its benefits, concerns, limitations as well as what steps organizations could follow in order to obtain maximum insights and benefits from SMA. Topics related to analysis tools, visualization and multilingual analysis have also been discussed with examples.

Chapter 3: Research Methodology – This chapter in detail explains what types of research methodologies have been adopted for this study. Details about target audience, number of participants, methods involved, data collected and research analysis has been included in the chapter.

Chapter 4: Research Results – Data obtained from the various research methods have been analyzed and the findings have been discussed in detail in this chapter with the help of relevant figures and charts. Hypothesis
testing results have also been explained and discussed in this chapter. A framework showing how various sectors within the industry can benefit from SMA along with relevant studies and examples also form part of this chapter.

Chapter 5: Evaluation – Critical evaluation of the study, methods adopted for this research as well as the results obtained. Limitations with respect to the research and how future work can be built upon this study have also been discussed in this chapter.

Chapter 6: Conclusion – Conclusions including lessons learnt from this research will be explained in this final chapter.

The reference list consists of all references used in various chapters of the report in IEEE format. Appendix section contains:

Appendix A consisting of survey questions and relation information and consent form.

Appendix B consists of the poll screenshots.

Appendix C consists of IBM – SPSS screenshots which were used the hypothesis testing for this research.

Appendix D consists of information and consent form for the focus group discussion.

Appendix E consists of R Programming codes and screenshots.

Appendix F contains of images as needed for some sections.
2.1 Social Media Analytics (SMA)

In today’s business world it is known that marketing and sales is impossible without data or information. With Big Data Analytics as the major topic of discussion and social media being one of its large source of data, the importance of its analytics and benefits can be justified. Zeng et al. defines SMA as, “Social media analytics is concerned with developing and evaluating informatics tools and frameworks to collect, monitor, analyze, summarize, and visualize social media data, usually driven by specific requirements from a target application” [1]. This research on social media and its analytics explains the concept of the kind of information generated and that can be obtained from it along with emphasis on its analytics and benefits to gain valuable insights. With the increase in the use of smart phones, we live in an era where social networking is just a tap away. Instant updates of facebook posts, twitter tweets, Instagram shares gives live feeds of things happening all around the world, where the barriers of time and distance is completely eliminated. Each of these posts contains different views, opinions and comments about various things from people from all over the world, breaking all barriers of time, distance and space and literally connecting the world with everyone and everything. The data generated, with the right tools can be analyzed to understand new requirements, the changing trends and realizing what the world requires. Generation of data keeps increasing with time, but bulk data is useless, what today’s world needs is to convert the massive amount of data generated to create a big impact. Analytics is about gathering, cleaning and preparing the right amount of data and converting it into information which becomes knowledge. This knowledge is used to answer questions related to business needs, academics, health, science and is the key towards valid decision making. Social Media Analytics (SMA) is also commonly referred to as Social Network Analysis (SNA), a better emphasis on the topic is shed by one of the world’s leading IT research and advisory company, Gartner. SNA as defined by Gartner are tools that are used to identify work patterns and similar structures between various groups of people, individuals or organizations [2]. These patterns or relationships are identified by collecting and mining data from multiple social media networks and helps determine the current trends, helps in targeting potential customers as well as shows information flow among people.

2.1.1 Terminologies related to SMA

Some important keywords one must understanding before reading any article related to SMA are mentioned as below:
**Text Mining** – Also referred to as textual analysis or knowledge discovery; is the study of text in order to find similar patterns and characteristics in orders to understand population, culture and trends. Tan [3] understands a text to be a most natural form of storing information. Text mining is performed in order to extract new and interesting information in order to understand a certain sample of the population in a better manner.

**Sentiment Analysis** – Information gathering technique in order to understand the mood or opinion of the public over certain topics. Pang and Lee in their study emphasize sentiment analysis as a mode of information retrieval by performing computational operations on sentiments, opinions and subjectivity in any text [4]. Depending on the fields of usage it is also referred to as opinion mining, subjectivity analysis, review mining and appraisal extraction.

**Data cleaning** – The process of removing or correcting corrupt, duplicate, incorrect, unformatted data in a dataset on order to produce better quality analysis and results. Apart from manually checking datasheets for such errors, powerful software using strong algorithms are used to identify and rectify similar errors or issues. Also known as data cleansing or data scrubbing.

**Data protection** – The data cleaned and formatted must be stored and protected according to The Data Protection Act (1988) standards. The data obtained could be from various sources and could contain personal information of individuals as well as corporate data of businesses or government data. In order to avoid legal hassles, protection of data is a crucial necessity within this area.

**Data Analytics** – Analytics performed on data obtained from social media or databases using powerful and sophisticated tools in order to derive required information and to study changing trends and patterns.

**Data Visualization** – One of the most important aspect of any sort of data analysis is the visualization of the information gathered. Not everybody understand figures and numbers and therefore it is essential to put forward the information obtained in form of images or figures that is understood in general in order to enable effective communication.

### 2.1.2 Top Grossing Social Media Platforms and their leveraging

According to a survey conducted in September 2014, by the PEW Research Center it was found that Facebook which celebrated its 11th year recently on February 4th, 2015 remains the most popular social media network continuously for the last three years [5]. Facebook is followed by LinkedIn, Twitter, Pinterest and Instagram; and the number of users grows rapidly every year. With all these platforms with millions of active users, the SMA trend is no longer just a necessity but has become the vital core of merging marketing, IT and communications infrastructure to derive informations and gain insights into various topics and areas of research.
A number of powerful analytics tools have been buzzing in the market, before exploring the tools it is essential to understand what form of information can be obtained by leveraging these social networks. Table 1 below gives a brief overview of the aforementioned top five platforms and what kind of data can be obtained from these platforms in order to obtain actionable insights. The information provided in the table has been obtained from a number of valid online sources as of August 2015.

<table>
<thead>
<tr>
<th>Platform</th>
<th>Facebook</th>
<th>LinkedIn</th>
<th>Twitter</th>
<th>Instagram</th>
<th>Pinterest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview</td>
<td>Social networking platform where users can share, like, update and receive notification on various posts.</td>
<td>Professional or business-oriented networking platform. Users can make and build new connection to improve their networks.</td>
<td>Social networking platform that allows its users to post 140 character posts called as tweets.</td>
<td>Pictures/Video sharing, social networking platform. Hashtagging became popular because of Instagram after Twitter.</td>
<td>Online mass photo sharing platform. Users can post and pin pictures posts they like. This platform is more popular among the women according PEW Research study [5].</td>
</tr>
<tr>
<td>Audience strength</td>
<td>1.44 billion</td>
<td>364 million</td>
<td>302 million (Active users)</td>
<td>300 million</td>
<td>47.1 million</td>
</tr>
<tr>
<td>Languages supported</td>
<td>70 languages</td>
<td>24 languages</td>
<td>28 languages</td>
<td>25 languages</td>
<td>22 languages</td>
</tr>
<tr>
<td>Information that can be obtained</td>
<td>• Target potential customers, suppliers, and partners; gain information related to them. • Obtain positive and negative feedbacks. • Brand popularity among public by analyzing likes and shares. • Finding competitor details and current trends. Note: Unlike Twitter and other social networks, Facebook networks are not publically available and requires user password.</td>
<td>• Build networks with new clients, suppliers and partners hence reestablishing the business. • Find skilled labor or employees. • Gain viral contents and effective or trendy headlines or topics. • Analyze online polls to gain statistics over certain topics.</td>
<td>• Analyze the hashtags on various topics to understand public’s sentiments or opinions on the topic. • Find out the most trending topics or headlines. • Research can be performed on twitter data that can be streamed easily via the API’s.</td>
<td>• Understand brand popularity by number of likes and comments obtained. • Find competitor information and marketing strategies on social networks. • Comments and opinions on brands and products to enhance customer experience. • Obtain live feeds regarding business events and venues.</td>
<td>• Explore more about the brands audience and find out common interests. • Obtain pin insights via the number of pins, impressions and clicks. • Information such as gender, location and interests of your audience. • Currents trends on the network. Note: Pinterest only recently made its API open for application and research purposes.</td>
</tr>
</tbody>
</table>
2.2 Analysis and visualization tools that can be used for SMA

Various type of information can be obtained through social media platforms by engaging in different types of analysis, the data analyzed can be also visualized in a number of creative ways. Making it easy to communicate and share with other people and also infusing better understanding among the viewers. Few trendy visualizations that can be created using social media data and tools are explained in the following sections.

2.2.1 Visualizing brand activity and personality

Brand activities and engagement on various social media platforms can be analyzed and viewed using analytical tools such as Brandwatch in order to understand consumer engagement levels, the different genders, new trends and topics on the pages. Figure 1 shows an image of a brand activity dashboard made using Brandwatch. This is only one of the example, there are tools such Hootsuite, Tableau and many more that perform similar analytics and meaningful details visualizations. These images can also show top influencers, top tweets, tops posts and demographics of consumers and fans. The social buzz created by the brand can be realized by viewing the number of posts, tweets or shares related to the page, the number of time the brands hashtags are used, find out the top influencers on these networks and the frequency at which the users are active on these platforms.

![Figure 1- Brandwatch dashboard showing different social media activities of a particular brand [6]](image-url)
Heat maps are another effective mode of visualizing brand activities since they map users, fans and followers according to their geographical location and shows them clearly with different color schemes on maps. This helps analysts understand locations where brands are quite popular and the areas where they need to improve upon. Figure 2 shows a heat map representation using Twitter Maps after Beyonce’s (Singer) new album was released. The map shows user activity and location based trends depending on tweets. Other software such as heatmap and Open Heatmap also perform similar analytics and visualizations not only to find user location but also shows heat maps on texts depending on various sentiments and opinions as well as produce heatmaps on spreadsheets. Google Analytics, RegioGraph and MapsData are among the tools which can track customer location in order to plan sales according to territories and various locations.

2.2.2 Sentiments and opinion visualization
Opinions and sentiments over a particular topic or brand name helps companies realize what the consumers and the public have to say or voice out. This information sheds light over areas that could be improved, products and services that needs to stay, those that need to be eliminated, user requirements as well as reviews on products services. Figure 3 shows Sentiment visualization related to hashtag #UAE that was created using a free online software called Sentiment Viz [8]. It gathers twitter data related to the keyword and arranges them according to their sentiments on a graph [9].
The sentiments similarly could be depicted over a **Timeline** using the same free online software to show various sentiments over a period of time. The sample result obtained for the keyword #Etisalat is shown in Figure 4. Etisalat is United Arab Emirate’s major telecommunication service provider and serves over 6.3 million mobile subscribers within the U.A.E.

**Word clouds** are also an effective visualization technique in order to view sentiments and opinions related to a particular keyword or topic. Figure 5 shows various sentiments collected over 2014-15 by the US Government on the changing gas prices in the country. Sentiment word clouds show the most frequent occurring words in larger sizes and least occurring sentiments in smaller sizes. Word clouds don’t necessarily have to be associated with only sentiments or opinions, associated topics or related words can also be viewed using word clouds. Word clouds can be created using data obtained from twitter using R Language, Wordle online software, SAS visual analytics and other tools.
2.2.3 Common other visual representations

LinkedIn connections can be visualized using **cluster and node visualizations**. Figure 6 shows LinkedIn connection of own account (Chantelle Stephen). An enlarged image will be provided in Appendix F. The visualization was created using online free social tool called Socilab [11], where the user can sign in using his/her LinkedIn account. If viewed closer, different nodes representing different sectors of industries can be as well as clicking on these nodes can highlight mutual connections. Such information when used on company profiles can help locate or find suitable suppliers, clients or influencers who may directly or indirectly may be able to benefit the business. IBM ManyEyes, MyWebCareer and InMaps are among the multipurpose tools that can be used to leverage and gain insights on LinkedIn network.

**Bubble chart** visualizations are also a commonly used visualization technique to unleash insights related to a particular topic. These types of charts show various entities according to their dimensions or popularity based
on the topic. For example, Figure 7 shows a bubble chart showing popular digital camera models from data obtained through Flicker website. The visualization was created using IBM’s ManyEyes, an advanced data visualization software with various charts and visualizations to choose from.

The above mentioned tools and visualization patterns and only the most commonly used types, a number of various other techniques and tools are available and used in the market. The ability to select and utilize the most effective tools could provide detailed understanding and insights over various topics that could help increase business values. Tools and methods use may vary according to business needs and social media strategies as companies could have varied demands and requirements from various social media platforms.

2.3 Benefits of Social Media Analysis

With clients being engaged and building relationships with brands and services online, keeping updated with the current trend of social data revolution emerges to be vital in order to maintain customers and to gain their loyalty. Adopting SMA strategies are known to show the following customer, business, marketing and management related benefits and can be justified according to previous studies and research.

2.3.1 Customer related benefits

Customers or consumers are every organizations number one priority and are vital for the company’s success. Social media has proved to be very effective in engaging with customers and understanding their requirements. Customer related benefits directly or indirectly pave ways to all other benefits such as business, marketing and management and are hence of prime concern. Areas where SMA shows brilliant customer related benefits are:

- **Faster and increased customer engagement** – Relationship marketing is one of the most traditional form of marketing being practiced for centauries where customer loyalty is obtained by building trust relationships with client by proper customer-engagement methods. Social networks provide real-time interaction interfaces between consumers and organizations so as to connect and build better relationships with them [13]. This medium also enables better customer service engagement as well in a much quicker manner. Sashi believes that customer engagement follows a digital and non-digital mix cycle of connection, interaction, satisfaction, retention, commitment, advocacy, and engagement [14]. Keeping this cycle and adopting proper customer engagement cycles will facilitate increased customer loyalty and retention.

- **Understanding clients in a more efficient and in-depth manner** – Better client understanding can be obtained through SMA which in return will help predict customer behavioral patterns as well as determine current needs. Linguistic analysis can be employed on social media platforms in order to analyze user generated content to realize current trends and requirements in the market [15]. Natural Language
Processing (NLP) is used as the basis for linguistic analysis in order to find patterns especially on large twitter corpus, which is a dataset of more than a million tweets. Consumer behavior on individual as well as group level can also be analyzed and predicted using similar techniques. Asur and Hubermans [16] study on predicting future with social media data has shown results and proved that data obtained from social networks such as Twitter can be manipulated to forecast future outcomes using sentiment analysis on the texts. Zhang and Pennacchiotti supports the above study by also proving that even purchasing behaviors of consumers can be predicted using similar techniques which was experimented on Facebook-Ebay datasets [17]. The study shows the relationship of consumer’s interaction on social media platforms and their purchasing patterns on e-commerce websites such as Ebay.

- **Learning customer sentiments and opinions about products and services** – Opinions expressed on social media platforms play a major role in understanding what the public’s opinions are regarding a certain brand product, solution or service. Most of the feedbacks whether positive or negative is expressed over social networks, blogs, forums and other digital sources. Sentiments can be positive, negative or neutral and depending on their percentage or count, companies can realize what needs to be changed or improved in the products or services so as to gain acceptance. Jansen et al. study on microblogs show that around 20% of these networks show mention of brand names [18]. Consumers tend to engage on social media so as to gain attention from companies as well as their connections and this action must be taken seriously as it created a huge impact on the brand whether in a good way or the other.

### 2.3.2 Business related benefits

Business related benefits are related to the brand awareness and financial benefits of a company. If social media networks can be efficiently leveraged and customers are kept satisfied, they lead to the following benefits:

- **Increase in revenue and profits** – Gaining fans and customer loyalty by proper utilization of social media strategies in return provides good annual revenue as well as improved profits. Barry C. et al. research on employing social media shows that those consumers who interact with companies over the social media are 40% more likely to be loyal than the others [19]. According to a Harvard study social media engagement on platforms such as Facebook and Pinterest could lead to an increase in store sales even [20]. The study shows that one in three Facebook users buy a product after liking, sharing or commenting on an advertisement. Social media adopters have seen a major differences in their business turn overs when compared to those who stick to the traditional methods.

- **Measure ROI using Attribution Modelling** – First introduced by Google Analytics, attribution modelling helps confront the concerns regarding ROI measurement to a certain limit [21]. Attribution modelling is an analytical technique that can tell how or what paths did a consumer follow to reach a particular
advertisement or website. Most of the information is obtained by tracing mouse clicks and search events that lead the user to perform the desired outcome such as purchasing a product. This has proved to be one of the biggest analytical breakthroughs by Google Inc. ROI’s could be measured using this technique by calculating number of clicks, user views, events searched for, advertisement clicks and impressions along with related online and in store purchases. There are single channel and multi-channel attribution techniques [22]. Single channel events includes measuring one particular event such as a mouse click or advertisement view. Whereas multi-channel refers to effectively measuring more than one event contributing to the desired outcome. Sequence, frequency, path, duration, and occurrence are the main factors that help attribution analytics. One of the ways organizations use to track user online activity is through loyalty programs, where customers tie their email-identity to online activities. Though this process requires time and effort, it yields profitable results.

- **Identify potential influencers** – Influencers can be consumers, clients, suppliers or anybody who helps support or drive the flow of business. Brands do not necessarily focus on the entire population or all types of buyers, but instead has particular targets in mind and is referred to as influencer marketing which aims at potential consumers or suppliers. Social networking has laid base as the perfect platform to reach out and to build relationships with these people. Booth and Matic’s research on mapping and leveraging influencers using social media justifies the above statement [23]. The study shows that understanding the engagement between influencers and brands through the social media could provide results that can be measured and will contribute towards the brand awareness and integrity.

### 2.3.3 Organizational related benefits

Benefits that affect the organization or the management, directly or indirectly are related to customer and financial benefits as those listed above and could lead to the following benefits:

- **Successful marketing strategies** – The efficiency of various marketing strategies employed on social networks can be realized by comparing and contrasting with competitor strategies. Also new marketing trends can be realized, leading to ideas in order create customized campaigns and promotions resulting in strong and solid based marketing strategies. The hash tagging phenomenon became widely popular as soon as people got to know that any keyword preceded by a hashtag (#) makes the information lawfully available to the public. This is a major marketing strategy example, that’s shows how such a movement became viral through a social network like Twitter [24]. Companies then started to gather data through Twitter metadata and would search for the particular keywords to perform analysis and gain insights.

- **Innovative minds can be employed** – Analysis of professional social networks such as LinkedIn where business and social media meet; would help connect with skilled labors with high expertise and knowledge
in various areas. It is a common phenomenon in the corporate world where employers try and gain connections with highly efficient employees through such social networks. Fresh talent as well as business networking can also be found and engaged using such platforms. Map visualization on such websites provide information about connection strengths and related searches. Also knowledge about what jobs are being commonly searched for or what are the common job trends within the market can be analyzed.

- **Improved production and supply chain planning** – “Being tuned in to changing customer tastes and behavior, businesses can anticipate significant changes in demand and adjust accordingly by ramping production up or down” [25]. Information obtained from social networks helps understand consumer needs and what kind of features they are looking for from products or what facilities to expect from services. Mithas et al. study shows that the analysis of such information helps understand production demands and hence facilitates for flexible production and supply chain planning [26]. Such planning results in reduced costs as it is planned well ahead with reduced marketing time and also leads to improved sales. Social media analytics could help analysts understand consumer as well as their changing demands.

### 2.4 Concerns related to employ Social Media Platforms

Organizations that use as well as those who do not engage in social media activity and its analysis have different types of concerns depending on their level of understanding and experience within the area. Some of the major concerns are listed in the following sections.

#### 2.4.1 User related concerns

- **Online deception** – Deception on social media is quite a common crime in today’s world, there are also advanced cyber-crime prevention and anti-fraud detection centers to counter such crimes. Tsikerdekis and Zeadally’s study on “Online Deception on Social Media” reflects on how deception prevails within the social media platforms; related to content, users, communication channels and even in a combination of all three [27]. The study shows that even though the distance between the deceiver and victim might be large, the damage done can be massive. Organizations and individuals must protect themselves with software solutions that overcome the traditional spaces and techniques that maybe be prone to such attacks and must use necessary precautions. Since, user identity and information can be easily tampered on social media, it poses as a hot spot for various forms of cyber-crimes and majority of the times the victims don’t even realize it until it is too late.

- **Privacy violations** – The public’s perception of posting on social websites have changed due to the increased number of identify thefts, frauds and online deception. People are more focused in posting in a private environment among their circle of friends and family. According to Jason Schwartz, the cofounder
of IPhone app called “Matchbook” the age of “Social sharing” has come to an end [28]. He refers to people wanting a more private life and not keen on exposing all their view to the entire world. His studies underline his opinions where he believes that the current trend now follows a more “classier” approach. Because an individual or a community does not care about public opinions until it is a public movement, for example one wouldn’t want to share their private pictures with the entire web but would choose to with their close circle of friends. Which is why social media giants such as facebook, twitter and Instagram have included more privacy options in their settings now so that their users can target a specific group of people according to their liking. Hence the coming years will be more focused on private group sharing and can be renamed according as an Era of Social Circles. People’s hesitance on sharing views and opinions on social media or in public makes it difficult to gather enough amount of metadata to conduct analysis. This might result in poor accuracy of the results due to lack of data volume and velocity.

2.4.2 Business related concerns

- **Data integrity issues** – Validation of data sources is one of the major issues while performing analysis on any form of data. The sources of reviews, forums, blogs and other digital sources can only be trusted to a certain limit because those are not academic, validated or published data. Due to the different types of users of different levels of education and knowledge, the content varies from excellent academic work to bogus or spam. Information obtained from various sources related to various field for analysis could be incorrect due to the quality of information obtained thus affecting the quality of the results. In Gayo-Avello’s study where he plays the Devil’s advocate while mining and extraction of information from twitter in order to show prediction outcomes of the US presidential campaign in 2008-2009 [29]. This research showed that the data obtained via the tweets could not help predict the correct potential Presidential candidate by performing sentiment or content analysis. The problem was not only the data collected but also how to ignore biased data contained in these social networks. The study concludes that until majority of the population uses the social media for voting purposes this analysis would remain incorrect and inconsistent. The data validation issue poses as one of the major pitfalls when it comes to information analysis and is taken seriously. Agichtein et al. research on mining high quality information from the social media shows various perspectives in order to obtain high quality content for analysis [30]. The study helps analysts face the data-validity issues and proposes techniques to check if the data being analyzed is of high quality and would provide accurate issues.

- **Legal consequences** – With various types of issues in place along with privacy violations, legal concerns comes into play and are necessary. Laws and policies must be adopted by organizations due to the increased use of electronic devices by employees. Jennings et al. research [31] studies clearly shows 100% of social
media usage among employees during office hours, among the top social network platforms are Facebook, Twitter and LinkedIn. Organizations use social media mainly for recruiting, marketing, networking and analysis purposes and all such networks are open to the employees as well. Employers cannot curb their employees from voicing their opinions on such networks since, public speech is protected by employee labor laws. Employees may share sensitive corporate information over such networks which maybe be susceptible to ill use. To curb such practices, employers can ban usage of social media within work environment by adopting certain policies. Hence, leakage of confidential information and trade secrets can be minimized. By adding such social media policies, employees are not tempted to breaking federal security laws and employers don’t have to be concerned and waste time on such legalities. Majority of the organizations are aware of such issues and have adopted security policies whether for company devices or bring your own devices (BYOD) in order to track employee actions. For example, IBM’s Maas360 is used for Mobile Device Management (MDM) and many other solutions as such are available in the market. Legal issues are also concerned with user data protection issues, if gathered data is not encrypted or protected in a suitable manner, it can lead to heavy penalties and legal consequences.

- **Concerns related to Return of Investment (ROI)** - One of the major concerns of corporations is return on investment (ROI), management might be worried that investments will go futile without proper knowledge or skill in order to handle and analyze information obtained from social media data platforms. Business can create values such as co-creation with users, marketing communications, understanding the trendy demands, find value of social commerce from these analytics. McKinsey Global Institute studies estimates that adopting social technology can add an additional few billions in almost any sector, for example his studies proved that social technologies could additionally make a $150-175 billion annual revenue within the automotive industry [32]. Production planning plays an important role within the automotive industry, collaboration of social media analytics along with research and development can lead to great success and this lead to such huge returns. The right way to conduct social media monitoring or analysis is to choose the right platforms, to know which social networks are trending according to location and language, and the second important aspect is to choose the right keywords to perform any sort of analysis. The common problem is that organizations look at obtaining immediate returns or expect an immediate shoot in their profit levels. In general it is not realized that, in order to earn some form of business value it the consumers or the audience who must benefit from them first, and this happens only after the initial momentum is created and take time. Hoffman et al. study supports this statement where his findings show that instead of measuring ROI’s using traditional methods, consumers motivation and their investments when engaging with the company’s brand must first be realized [33]. The study also conclude that engaging with social media does not provide short term results, instead it is a long term relationship with customers and is built on mutual trust.
Prejudices related to SMA – It is common among the public to doubt new technological practices or to make immediate misconceptions that it will be too difficult to even bother. Companies may fear that they do not have the right amount of skilled labor or knowledge without even making an effort. Another common misconception is that social media and its analytics are only reserved for the marketing team or the Chief Marketing Officer (CMO) in an organization. Social media strategies should involve all level of employees and should not only be used as a medium for marketing but also as a communication channel. Where customer service agents, sales agents, business analysts are equally responsible as the marketing team to reach out to the consumers and their requirements to keep up the brand name and loyalty.

2.5 Limitations of Social Media Analysis

Apart from the benefits and concerns of social media analysis, being still a major area of research and development, it is known to have a number of limitations:

- **Data quality affecting analytics** – The quality of results are deteriorated due to the poor quality of metadata or information gathered for analysis purpose. The importance of validation of data sources has been explained in section 2.4.2 and is one of the major concerns related to social media analytics. Due to drastic variety and huge volume of user generated data it is very difficult to validate each and every source but is not impossible. Recent studies of Todwal and Wanjari on community blogs such as Yahoo! Answers shows promising results where high quality content is obtained avoiding ambiguous content like connection details and quality ratings [34]. Quality of social media data for analytics can be estimated by link analysis or linked based algorithms as cited in this study, this ensure the quality of the results as well. On the other hand Dewan P. explains why the data quality of social media platforms such as facebook is minimal and of bad quality [35]. The study states that 72% of facebook users set their posts to private whereby it is not accessed via API streaming for analysis purposes. Twitter feeds on the other hand have be a much favorable social network used for research and analysis since it is easy to obtain tweets by searching by the keywords or hashtags. Again these keywords also bring those tweets which are of no use or contain malicious information and data cleansing tools must be employed to get rid of these which maybe expensive and time consuming. On the bright side, most of the social media analytics tools contain powerful data cleansing and preparing tools which does automatic procedures and data cleansing techniques which provide good analytics.

- **Limitations of sentiment and linguistic analysis** – Sentiment analysis especially automated sentiment analysis performs poorly on biased data. For example sarcastic comments and exclamations maybe not be analyzed as positive, negative or neutral in the right manner. The major issue concerned with sentiment analysis is the order of the words, a sentence like “Thank you, but I hated it”, maybe be classified as a
positive sentiment since the word “thank you” occurred first. Also since sentiment are expressed in varying manner from person to person using various slangs, expressions and sometimes even emoticons it is difficult for the sentiment dictionaries to accommodate all such data and to polarize them correctly. In defense, the area is still improving with time and is still under research and there are a lot of studies and experimenting happening within this topic. Hogenboom et al. study on exploiting emoticons and polarizing sentiments accordingly shows a promising part in the field especially with the increased usage of emoticons [36]. The research was based on the fact that in most cases the emoticons represents the entire sentiment of the sentence and hence can exploited and used in lexicon-based sentiment analysis to provide results.

- **Loss of revenue due to poor social media strategies** – Poor knowledge of leveraging social media platforms and improper background study and research might lead to major social media failure. No form of analytics builds its own strategy, instead the strategy is made by the organization and analytics are used to produce results by following the strategy plans in an organized manner. Aral et al. study of social media, society and business it was found that changing communication and consumer trends must be adopted by organizations in order to keep abreast within the market [37]. His studies also showed that medium to small scale organization lack guidance about social media and strategies to benefit the most from it. Research conducted by Fisher and Reuber compliments the previous study showing that though entrepreneurs develop effective communications and relationships with consumers; lack of knowledge on how to engage effectively shows a gap in accomplishing goals [38]. The study points out social media shouldn’t just be used as a medium of communication but also entrepreneurs must learn how to leverage such social networks and how they can maximize benefits. For example one of the famous social media controversies on strategy was of Kenneth Cole, when in 2013, during the Syrian civil war, the American designer decided to take light of the situation and posted a tweet that caused great uproar and controversy [39]. The designer’s actions are plan of a marketing strategy commonly termed as “controversial marketing”, though not favored by many, and has a history for controversial tweet like the one during the Egyptian protests in 2011 [40]. Though the designer has publically apologized for insensitivity in both situations, he claims he was well aware of the consequences and the marketing strategy has indeed bought him brand publicity. This example shows that strategies are not rules that can be applied as a uniform in all sorts of business, but in fact it varies from one organization to another. Proper consumer and background study with knowledge about the field with result in effective means to gain revenue.

- **Not a silver bullet to turn customer centric** – Relying on social media solely cannot make any organization or enterprise customer centric. Recent marketing research shows that there are three types of social media users who are categorized by the time they spent on social media, their actions on these platforms and their post frequencies [41]. They are called as enthusiasts, lurkers and dabblers in order of their social media activities in a decreasing manner. The study shows that 85% of updates and interactions
come from enthusiasts who only comprise of 28% of the population. The other two groups which is the lurkers and the dabblers have no interest in being part of such social interactions. Whereas these two groups of people must be loyal customers who directly purchase through the stores. The study shows that the behavior of all these groups varies from each other and therefore analytics obtained just from one group of people causes gaps in the analytics. Therefore other methods must be adopted to reach out to the various types of people, relying purely on social media platforms alone won’t do the trick.

- **Immediate ROI measurements are not possible** – Social media analysts often find it difficult to their management or organizations about what gains they are benefiting from employing social media. What is not realized is that it is a step by step that take time, investment and effort before obtaining gains. It is not always about ROI’s, what organizations must focus on to learn the craft of SMA properly and to understand how their customers are benefiting from it. Social media benefits are generally difficult to measure, factors such as online traffics and fan count can be measure easily but those like tracking fans, likes, popularity does not necessarily have to give ROI figures on charts. Leaving traditional ROI measurement aside companies can still use certain parameters to understand benefits of social media. For example, Vitabiotics, a British company selling multivitamins and health products all around the world engaged on social media platforms for community feedback [42]. Results showed that annually they had hundred thousands of profit when compared to traditional marketing patterns. Blanchards book [43] and Hoffman et al. research [33] on measuring social media ROI, together help conclude with the explanation above.
2.6 Multi-lingual analysis – A critical requirement within Social Media Analytics

With over billions of adults engaging on the Internet from all the world; lingual diversity is at its maximum in this augmented area. Bloggers, user journals, academics, articles, reviews all come in a variety of languages. In general, users are normally comfortable using native languages than foreign languages as they feel they can express their opinions in a more creative as well as open manner. English is used all around the world as a common language for communication between individuals, and there are a number of techniques for mining text in English. The entire world being involved with the technological changes and with the social media, has caused increase in the use of multilingual languages. There are number of multilingual keyboards which allow users to type from laptops, smartphones and similar electronic devices in a number of languages. Social media mining or sentiment analysis is easily possible in English as there a variety of tools that have been developed and improved over the years. When analysts understood the trend of people typing in native languages like Arabic, Spanish, Hindi (and other Indian Languages), Italian and other languages research began on how to conducts multilingual social media or sentiment analysis. For example, in the United Arab Emirates where Arabic is the native language; opinions about food, leisure, fashion, business reviews and advertisements are a lot in Arabic and in such a geographical position where Arabic is widely used the requirement for Arabic sentiment analysis or opinion mining is important. In Figure 1, screenshot of an Instagram post from the official page of New Look Middle East. New Look is one of the trending fashion brands especially among high school and college students and is located in majority of the malls within the Middle East as well as abroad. As the image below shows, the advertisements are in English and Arabic, which proves that a large number of readers are Arabic speakers. The image also shows comments in Arabic from various users which is enlarged and pasted below in Figure 2. Not only English and Arabic are the languages used on social networks within this region, being a culturally diverse hub with more than 89% [45] of the population being expatriates; the number of languages used within the region and similarly on social media is plenty. Also with UAE being the host for the Expo 2020, the number of business partners, vendors, suppliers and consumers all around the world shows the critical need of employing multilingual analysis within the region.

![Figure 8 - New Look Middle East Instagram Post](image)
2.6.1 Relevant research and studies related to multilingual analysis

Previous studies and research shows that multilingual mining can be effectively employed with the correct techniques and technology. Hsieh et al. describes the importance of mining social networks for event alerts, they have aimed at Bilingual mining of English and Chinese while creating a Social Event Radar [46]. The main objective behind creating this innovative system was to track social networks and alert events that needed care and was mainly focused on Food safety and Brand Monitoring. Bi-Lingual sentiment opinion analysis (BSOA) and Social Network Influential Analysis were the main building blocks of these systems, the system was used show alert at public safety, enterprise as well as government level.

Elhawary M. and Elfeky M research on creating an Arabic mining system, focusing mainly on business reviews shows how different classifiers can be trained to identify and extract reviews in various languages [47]. The experimental system consists of two main components, a classifier to identify URL’s and label them as reviews, shop store and news. If reviews, the content in Arabic is identified by an extended Arabic sentiment analyzer. The system then polarizes the sentiment of words extracted based on certain scores which helps determine the sentiment involved in the whole review or sentence and the intensity of these sentiments. Although, the experiment was set up in a small scale, similar concept could be used to build large scale systems in order to mine data in other languages especially to mine business reviews. Similarly, Rushdi-Saleh M. et al. research on Arabic opinion mining using corpus based approach [48]. Due to the difficulty of finding corpus designed for implementing Arabic sentiment analysis, the research was to build such a system in order predict sentiment polarity called Opinion Corpus for Arabic(OCA). The corpus training data were obtained from a number of websites containing mainly movie reviews which was generated and validated using step by step techniques. The results of the experiments where concluded by using tenfold cross validation on n-gram models in order to test the accuracy of Support Vector Machine (SVM) and Naïve Bayes (NB). It was noted that SVM was slightly better in accuracy when compared to NB and hence it was concluded that using SVM on OCA proved to be more accurate. These experimental setups mentioned above can be used in various website containing

Figure 9: Enlarged view Arabic comments on New Look post shown in the image above. [44]

[wedad_994 @mony_anmed91 @asoomme161] مشينا ي بدنات ١٦١
[asoomme161] بلا ١٦١ قاذام اليوموم أن شهادك
[joudbadran] فل سمعت هل بدانت بالرحب؟@newlookmiddleeast
[jodyabot114] جهير هاي هيلز مقاس ٣٥
[22]
multilingual languages and can be used to extract and analyze useful data in order to improve business review in any industry whether related to food, movies, fashion, technology and other similar sectors.

Brooke et al. study where they try adapting resources provided for English sentiment analysis to Spanish [49]. Already existing English semantic orientation calculator (SOCAL) was adapted into the Spanish language, and was testing by constantly compared with other alternatives. The experiments showed that this conversion showed significant semantic loss and comes with cost related to performance and accuracy. But the study is concluded by profoundly stating that “there is value in pursuing the development of language-specific resources, notwithstanding new breakthroughs in machine translation.” All the studies mentioned shows the requirement of multilingual sentiment analysis and the need for advancement in this area.

2.7 Effective utilization of Social Media Analytics for organizations

The building blocks required for utilizing and leveraging social media platforms to gain valuable insights and to take actionable decisions are shown in the flow-diagram in Figure 10 below:

2.7.1 Background Research

Before fixing on any form of business strategy and efficient background research is essential, which will help determine current trends and how to proceed in the right directions. Similarly understanding trending social media patterns in the market is required to understand and learn how to gain maximum information from these platforms.

❖ **Existing social media actions** – Any form of previous engagement on social media platforms must be analyzed and measured, which will help understand popular social platforms for client-engagement. Analyzing existing platforms will provide metrics on current user interaction, fan and followers count which

---

Figure 10- Building blocks for effective utilization of social media networking and its analysis.
will help determine benchmarks to realize active or poor platforms and also determine what type of information can be obtained from these channels. This study will also help realize what changes need to be made for improving brand awareness as well as to identify better methods of communication strategies.

- **Audience intelligence** – Conducting an audit on clients and gathering information regarding consumer age, nationality, genders, buying patterns and interests will help determine and categorize different types of consumers. The information gathered would provide answers as to what type of buyers to target on social media platform and what kind of campaigns and promotions to employ in order to improve sales or to promote customer loyalty. For example, if targeting professional audience then utilizing platforms like LinkedIn would provide more useful. Whereas, if looking for the millennial audience platforms like Instagram and Facebook would provide more results. Li Ying et al. research [50] on data mining and audience intelligence shows the importance of primarily understanding user attention on a webpage and supports the statements above. Audience intelligence can be gathered via surveys, interviews and from audience data obtained from media sources such as newspapers and magazines, which will help understand the trending social media habits.

- **Learn and compare with competitors and industry leaders** – Compare existing strategies and plan of actions with competitors by studying more about their social media engagement and strategies adopted to resonate with their customers. This competitive analysis will also help learn new techniques or elements that can be adapted into the company’s existing strategies. Basic data gathering can be obtained from competitor websites, social media platforms and by engaging with common consumers. Well established organizations with strong social media strategies should be role models and their engagement techniques, promotions, campaigns and other social network engagement should act as a motivation factor as well as become a learning platform.

2.7.2 Aims and Objectives

Before engaging in any form of social media activities or even before building any form of business strategy it is essential to have aims and objective identified and stipulated beforehand. This will ensure that all activities where time and money is invested in is working towards the goals and producing the desired results.

- **Identify Smart goals** – Goals are the final targets that are expected to be attained while engaging in social media activities. These goals must be realistic and achievable and must have a fixed timeframe associated with it. The goals must be answers to what is being expected from these social media platforms or why is the organization investing in this method. At some point of time these goals must be also measurable in order
to understand how it is helping the business for example the number of website visitors from different parts of the world can be calculated and analyzed.

- **Define brand purpose** – The purpose of the brand must be well defined in order to estimate its impact on social media platforms. It could be either for customer loyalty and retention, to promote brand awareness, to improve sales, research and feedback mechanisms, to target influencers or for all of these [51]. The brand is being used on social media platforms to build relationships with the customers and must be ensured that the purpose underlines the brand personality and keeps up with the reputation.

- **Brainstorming potential strategies** – Keeping in mind the aim, objectives and goals a team of professionals must be chosen within the organization to choose a particular social media strategy aligned with the identified goals. While choosing participants, it must be ensured that marketing and sales personnel’s must not be picked alone, instead it is important to included members from as many departments as possible to obtain innovative ideas to build a strong strategy. The discussions must be carried out in an organized manner and various ideas obtained must be displayed to all in the form of sticky notes, white boards or any other form. While choosing suitable strategies, the potential risks and threats associated should also be identified and discussed upon.

### 2.7.3 Building on a strategy

After brainstorming and assessment of various strategies, one suitable strategy will be agreed upon within the organization to be chosen and this must be done carefully [52]. Once a basic outline or idea has been implemented within the organization, the strategy detail such as channel planning, content development, communication of the strategy to employees are or critical phases of this process.

- **Channel planning and content development** – Identify which channel to engage in most, depending on the type of targeted audience, the types of posts and campaigns which will create the desired impact needs to be determined. Information related to the audience is obtained from the intelligence study performed earlier which will help identify suitable means of communication and thus pave way for effective content planning [53]. At this phase the report planning is also conducted in order to identify report intervals, types of reports as well as report content.

- **Team building and effective communication** – A team must be selected to keep track of the social media metrics and success; along with an inspiring team leader. This team not only measures social media metrics but also ensures effective communication of the various social media plans and actions among the employees. It is essential that the employees themselves are aware and engaged in these social activities, which will help increase awareness and activity within the platforms. The employee awareness program on the topic need not be only through formal discussions and meetings but informal gathering will also increase
employee attention and engagement. When the entire organization is involved, every employee is responsible for the success and will also promote innovative ideas and activities.

2.7.4 Evaluate and measure

Any series of actions or strategies can be called successful, only when they show proper results that can be measured and proven to the organization management. While building on a strategy there are certain benchmarks, targets and baselines that are laid out in order to measure its success. Therefore, it is essential to choose proper tools and techniques to measure these metrics as well as effective reporting and visualization to communicate this analysis to the organization.

- **Select suitable measurement techniques and tools** [54] – Success is generally measured in two basic methods, quantitative and qualitative. Depending on the goals and types of social media engagement suitable qualitative or quantitative methods needs to be adopted. Qualitative methods are easy to measure and include calculating fan or follower count, user engagement rate, click through rates which helps in analyzing website traffic, audience timing and conversion metrics which includes download rates, sales transactions as well as email sign ups. Quantitative methods are a slightly more difficult to measure but involves influencer analysis, sentiment analysis as well as linguistic or text analysis or analyze conversations. All these methods are supported by a number of powerful tools of software such as Sysomos, Oracle Social Cloud, Viral heat and many more which are available in the market.

- **Reporting** – The analytics, figure and results obtained from the desired techniques and tools needs to be reported and presented in a particular format decided by the management or team. These reports must be generated within a fixed time or day intervals in order to analyze how it is affecting business and the consumers. Analysis of these reports maybe even shows idea of how to improve on the strategies, help make decisions to take actions, show area of improvement and the impact of these actions on the consumers.

- **Visualize results** – Reports can only be understood by those who have proper insights on what is happening with the social media strategies. Employees not involved and other level of managements may find it difficult to understand these reports and there for it is essential to visualize results in a manner so that it can be understood in general. This is an effective communication method as well and will help discover more ideas and tactics to leverage social media platforms effectively.
2.7.5 Take Actions

Based on analytics and results produced in the previous phase actions must be taken in order to change, improve, remove or adopt new elements with respect to the strategies.

- **Decision making based on analytics** – Analytics show various information such as consumer behavior, requirements, current trends, sentiment over various topics, feedbacks, online sales and purchase behavior. Information gathered must be analyzed shows actionable insights and decisions must be made accordingly [25]. For example understand customer requirements about a particular product can help with production and supply planning. Analytics could also show which platforms customers provide feedback and need support. Accordingly more customer service representatives can be allocated or employed to tend to these clients.

- **Adjust strategy according to insights obtained** – Understanding areas that are faring well or require improvement or changes can help adapting the strategy in order to fuel efficient social media manipulation. With time, engagement on these social platforms teaches new lessons and shows ideas based on the analysis. These new methods can be applied to the strategy as well.

2.8 Social Media Analytics in Action – Success Stories

Top Multi-National Companies all around the world employ Social Media Marketing as well as indulge in their analysis to gain actionable insights in order to improve business. These organizations have been able to realize their Return of Investment (ROI) via these social networks. Success stories of leader’s from sectors of fashion, technology and food chain have been discussed below.

**Boden, UK** [55] – J.P. Boden & Co Ltd. is a British retail company using multichannel approaches for their business. They are mainly prominent within the United Kingdom and United States of America. While introducing a new e-commerce channel within both the region, the organization needed to understand their ROI based on customer interactions and navigation within the website when compared to other channels. Therefore, using a behavioral analysis software such as IBM’s Digital Analytics for Social Media, the company was able to understand consumer behavioral patterns on the website by tracing their paths according to their navigation. This helped the organization realize, what the customers were really looking at and where they faced trouble while navigating through the website. This not only gave the company insights about customer behavior, product demand and price concerns but also helped the customer service team to understand what sort of questions should they be expected from their customers and how they can tend to them to improve their experience. As a result, the purchases increased by 160 percent, as well as showed increased customer loyalty and saw the number of repeated visits improve substantially.
Symantec, USA [56] – The Symantec Corporation is an American technology company that produces a number of software’s related mainly to security, storage and backup. It also happens to the parent company of the famous Norton products. Being a well-established organization and having customers all around the world, keeping satisfied customers is a large task at hand. As social media popularity increase most of the issues and problems were posted on social media channels such as twitter, with the increase in social media usage tracking and responding to customers got out of hand. Employing a power solution such as the Lithium Social Web (LSW) they have achieved quicker response time, easier to monitor unique cases as well as multi-language support of up to 12 languages. The response volume has increased by 160 percent resulting in delighted customers whose positive sentiments via feedbacks could also be tracked by the solution.

Coca-Cola, Australia [57] – Coca-Cola an American company selling soft drink started a campaign in Australia called Share a Coke back in 2011, where the brand name on the bottles were replaced by common names among the Millennials or general population. This went viral all over the world around 50 countries until it hit the USA market last year (2014). In order to understand the public’s opinion over this movement, Networked Insights was employed to understand the statistics. By tracing Twitter and Instagram posts, blog, forums and other sources with #ShareaCoke it was found that 45% of the consumers were favorable towards the movement, 50% remained neutral and only 5% showed negative sentiments. Their marketing strategy along with their analytics proved to be a massive success within the market.
Chapter 3 – Research methodology

3.1 Research problem

With social media platforms being one of the most common methods of communication and an open source for information gathering, how effectively are companies leveraging such networks to gain valuable insights of the current market and trends? This study was conducted within the United Arab Emirates (U.A.E) where social media involvement is active and thriving. With the country already being technologically advanced, understanding the role of social media and its analytics looked interesting. Being a country with cultural diversity in population and speaking a number of languages other than English, the need for multilingual analysis within the region also needed to be studied.

An initial background study within 20 companies of different industries and sectors surprisingly showed less social media engagement and analytics than expected. This issue was more common among small-medium businesses. The reasons behind hesitation to use such an important source of information needed to be identified and researched. Also solutions to the problems had to be found in order to propose to help increase business values within the region, especially since UAE is the host for the Expo 2020. The Expo 2020 is a six month major event that will attract millions of visitors and therefore means increased business especially in the hospitality, airlines, tourism, and food industry. According to Dubai Expo 2020 official website, 25 million visits are being expected with above 70% from overseas [58]. Apart from hosting such a major event Dubai is known for its diverse culture of expatriates with various interests and who are up to date with international trends. Companies within the region have consumers not only within the country and the Middle East but also internationally. Hence, social media is an effective medium connecting clients from all over the world. How can the organizations effectively utilize social media platforms to obtain insights which will help them make actionable decisions? This was the basic foundation of this research. Keeping this in mind suitable research methodologies and analysis techniques were adopted to gain a better understanding on this context. The current trends among large, medium and small business needed to be explored and studies for better quality results. As well as understanding participant’s knowledge about social media analysis will also lead to a number of findings which will help in this research.

3.2 Research questions

- How active are the people within the U.A.E on social media platforms, what kind of languages or posts do they see commonly?
• How often does the people within the region purchase products or services on these social platforms? What drives them to do so?
• What are the major reasons for social media engagement and its analysis adopted less within the United Arab Emirates especially within small-medium businesses?
• Those organizations who leverage social networks, how effectively do they do so?
• Does lack of awareness in the topic have a huge impact on companies’ hesitation of not involving in social media platforms?
• The need or significance of multilingual analysis within the region?
• How does various sectors of industry benefit from social media analysis in the U.A.E?

3.3 Methodology
The necessary information to understand and learn about social media analysis, its building blocks, benefits, limitations and concerns related to it was learnt from previous studies and researches and forms the literature review (Chapter 2) of this report. A mixed methods research was then conducted for the study consisting of quantitative and qualitative analysis. As Creswell’s book suggests a mixed methods methodology in necessary in those studies where one data source might not be sufficient and where the initial results obtained need to be justified by a more in-depth explanatory findings [59]. Similarly this research required two data sources one from the surveys and polls and a more details findings with the help of a few professionals by engaging in a focus discussion. There are various mixed methods designs and the most suitable design ‘Explanatory sequential design’, was chosen for this project. Initial results will be obtained from the quantitative analysis which in this study is in the form of surveys and polls. The findings will then be backed up using a quantitative approach such as a focus group discussion in this case in order to explore participant’s views more in depth and to hence conclude with better interpretations [60]. Figure 11 explains the research analysis approach in a diagrammatic manner:

![Diagram of research analysis approach](Figure 11: Mixed Research Analysis method proposed for the study)

The two stages of the explanatory sequential designs adopted for this research will be explained in detail in the following sections.
3.3.1 Stage 1: Quantitative Research using surveys and polls

Quantitative research involves understanding the opinions, trends, behaviors and attitude about a sample population in a measurable manner which is usually in the form of numbers or figures. According to Cohen et al. book, quantitative research adopted empirical formulas and techniques for social research [61]. Therefore in order to understand the behaviors and trends related to social media and analysis within the United Arab Emirates quantitative analysis was chosen in the form of surveys and polls and is explained in detail as follows:

**Surveys** - The targeted audience for the survey involved 105 mid and junior level management from various industries with different backgrounds such as marketing, sales, IT, finance, HR and other departments. According to different company sizes – large, medium and small; 35 participants were chosen for each group. Prior to the survey every participant was provided with a consent form stating no form of personal information would be collected or circulated in manner and that the information provided would be used for academic purpose and could be subjected to publications. The participants were accessed via emails and mobile applications where links were sent and the surveys were filled out online. Information form, consent form and images of the survey can be seen in Appendix A. It consisted of three phases:

Phase One – Helped understand the impact of social media in the U.A.E. The frequency at which the population engage on social media platforms, the type of posts commonly seen as well as the various languages they use or see on social websites. The purchasing power of the public and the factors associated could be learnt from this phase.

Phase Two – Depending on the organization sizes, involvement in social media and its analysis were learnt. This phase helped determine effectiveness of SMA in organization who use such techniques and reasons behind those companies who aren’t involved in it.

Phase Three – The last phase was to help identify the participant’s depth understanding of social media analysis, its importance and its associated benefits. This also showed the awareness of social media analysis employees had in various industries.

**Poll** - Polls with two questions was conducted in order to find what the public’s opinion were over two topics:

- Purposes social media could be leveraged for?
- In an organization, who is responsible for ‘Social-media engagement’?

The polls helped support the findings from the survey is discussed elaborately in Chapter 4. Screenshot of the polls conducted can be seen in Appendix B.
3.3.2 Stage 2: Qualitative Research using Focus Group Discussions

Qualitative analysis methods are used to support the findings of the quantitative results as well as to understand participant perspectives, explore their opinions or to observe their reactions to a particular topic. According to Hancock et al. qualitative research provides better understanding or broadens aspects about how things came to be in the real world [62]. To ensure a better and deeper understanding within this research, the second stage of the research involved a **Focus Group Discussion** and took place over a period of two days. There were 8 participants who were selected from five industries – Education, Food, Retail, Aviation and Technology. The participants had professional as well as academic knowledge regarding social media analysis and most of them were even actively involved in it. The discussion happened over two days keeping in mind participants work timings and personal routines. The participants were all noted 2 weeks in advance about both the meeting and the time allotted. All participants were provided with a consent form prior to the discussion stating their participation was part of a study and that no personal information would be revealed in any form. The consent form can be seen in Appendix D.

**Prior to the discussions** - The participants were asked to beforehand to conduct a background study on the topic even though they were experienced in the field, in order to ensure better thinking and innovative idea. Two weeks prior to the discussion, after the notices had been sent to the participant’s, analysis on Twitter data was done using R programming. The code for this analysis was obtained from R Bloggers and ThinkToStart and was adapted to obtain the required visualization [63] [64]. The experience helped realize how researchers gather data from social media platforms such as Twitter and how R program was a simple and effective tool to gain information such as sentiments, to build word clouds and heat maps. The visualization of this results was part of the second day’s agenda.

**Day 1** – The 60 minute discussion with 8 participants was quite a learning experience and innovative idea filled the room. Four major topics with 15 minutes allocated time were discussed upon:

- General concerns related to social media analysis within the United Arab Emirates.
- Eradicating the stereotype “SMA is only for the Marketing team”.
- The necessity of multi-lingual analysis within the region.
- SMA benefiting the upcoming Dubai Expo 2020.

**Day 2** – The two visualizations produced during the analysis performed using R Programming was printed and circulated among all the participants. Looking at the three images the participants were encouraged to voice out the pros and cons with respect to each image. The 30 minute discussion help identify answers related to data quality, sentiment accurateness, multilingual requirements and visualization quality. The sessions also includes participants opinions on issues and limitations related to Social media analysis.
3.4 Research Analysis Plan

The Quantitative data is analyzed using a deductive approach. According to Crossman, sociology expert deductive reasoning can be defined as a bottom-up approach where, various patterns in specific topics are researched upon leading to interpretations and theories [65]. A similar theory has been used to study and analyze this phase of the research. The data obtained is analyzed and visualized using Microsoft Word and Excel tools and is represented in the form of different types of charts and figures. The findings and conclusions drawn are explained along with the analysis visualizations in Chapter 4. A hypothesis test has also been performed using IBM SPSS and is also explained in Section 4.1.4. An independent hypothesis or a two-sampled t test was used for this study. A hypothesis test is in general conducted to find a statistical difference or significance from a sample population data in order to obtain an inference. A two-tailed hypothesis test is used to find if the mean of two different sample population is the same or has any significant results. The research problem need to be formulated into a hypothesis (alternative hypothesis) or $H_A$ which generally states that there is some significant difference between the two means of the sample. A null hypothesis needs to be also formulated which must be theoretically opposite of the alternative hypothesis; which in general must state that there is no significant difference between the two means and is denoted as $H_0$. In order for a hypothesis to test positive, $H_0$ must be nullified or rejected. Before performing, the test a significance level or alpha value needs to be selected, this is probability of making a wrong decision while testing. According to the findings of famous statistician R.A Fischer, the alpha value is commonly taken as 0.05 [66]. If the p value calculated is less the significance value (<0.05) then the null hypothesis can be rejected and the research problem can be tested as positive.

The Qualitative data involves experiences or opinions of individuals and therefore a heuristic approach is adopted for analyzing and evaluating the information obtained. Kleining and Wits study states that’s a qualitative heuristic approach can be applied with any topic where individuals and their opinions are involved [67]. Similarly, the various topic were discussed with open perceptions while obtaining information and new ideas which also acted as a support to the quantitative results. Participants voiced out different opinions on the various topics, the similarities were discovered and the observations are summarized and recorded in Chapter 4 as well.

From the studies, research and understanding of the topic a small framework on how SMA can benefit major industry sectors has been made. The framework was research, designed, tabulated and explained briefly along with related studies and local examples and is included in Chapter 4.
Chapter 4 – Research Results

4.1 Quantitative data analysis – Survey and polls

Data obtained from the quantitative research, which is from the surveys and polls are analyzed and explained in the sections below. As mentioned earlier, 105 participants within the U.A.E took the survey as well as participated in the polls. Participants were all above the age of 21 and consisted of junior to mid level management from different departments and were selected after ensuring they were aware of any form of social media practices within their organizations. For better understanding, the participants were further divided according to their company sizes such large, medium and small scale industries, and there were 35 in each category. The company sizes were categorized as per Table 2 below:

<table>
<thead>
<tr>
<th>Company size</th>
<th>Number of employees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>&gt;50</td>
</tr>
<tr>
<td>Medium</td>
<td>50-200</td>
</tr>
<tr>
<td>Large</td>
<td>&lt;200</td>
</tr>
</tbody>
</table>

Table 2: Companies classified according to their number of employees

4.1.1 Phase 1: Influence of Social media within the U.A.E

The first phase of the survey was to understand the current trend within the United Arab Emirates with respect to social media activities.

❖ Social media engagement within the U.A.E

The results obtained depicted in Figure 12 shows that 53.98% of the participants engage on social media platforms on a daily basis, 31.86% engaging every hour and 12.39% checking every few minutes. The rest 1.77% which is a very small percentage and they check weekly at least. Options of “Monthly” and “Never” was not selected by the sample population surveyed. From this data representation it can be explained that residents within the U.A.E show high social media involvement where they spend ample amount of time on various social platforms, engaging in one manner or the other.

Figure 12- Pie-Donut chart depicting the social media engagement within the U.A.E
Types of posts commonly seen on social media platforms

To analyze what sort of posts participants generally see on social media platforms they were allowed to select multiple options of the posts they see on a regular basis. The data obtained was then analyzed and visualized using a column chart, which is shown in Figure 13. The maximum number of hits were for posts or reviews by friends, followed by advertisements, pictures, videos and shared images.

Languages other than English spotted on social websites in the region

In order to understand if multi-lingual analysis was a real necessity within the region, the occurrences of various languages needed to be determined. Participants were asked to select languages other than English they commonly saw on social media platforms. Multiple options were provided and participants were allowed to choose more than one option. The data obtained was analyzed over a radar chart to see the different variations between the languages and is depicted in Figure 14 as shown. As per the data obtained and from the chart, Arabic posts are
commonly seen over the social media, this could be in the form of advertisements, reviews, blogs, comments or in a similar manner. Indian languages such as Hindi, Tamil and Malayalam are also other common languages seen on social media especially within this region, where the Indian expatriate population is quite large. European languages followed with a slightly lower count and include languages such as French, German, and Spanish and so on. Other languages as per the chart were also predominant on these social media platforms.

❖ **Purchasing products seen on social media platforms.**

In order to understand how often people purchase products they find interesting on social platforms participants were asked to select from four options. Out of all participants 25.71% said they purchase very often, 38.1% said often, 30.48% said sometimes and only 5.71% choose the option “never”, as shown in Figure 15. This response shows more than 50% of the participants purchase often or very often on social platforms. This shows effective marketing on social networks could lead to customers buying products either online or the same from department stores.

❖ **Posts that convince participants to buy or invest in products or services**

Participants were asked to select one type of post that could convince them the most to buy or invest in any product or service. 50.48% voted for professional or peer reviews, 22.86% for attractive sales offers, 20.95% for brand advertisements and 2.45% choose others, as shown in Figure 16. From this data the influence of peer reviews or professional reviews on people within a network can be realized. Also attractive offers by brands can increase sales potential and draw customers to buying products and services.
Findings and discussions based on Phase 1 data gathering and analysis.

From the data collection and analysis of the first phase of the survey, the influence of social media is quite obvious. U.A.E being one of the biggest hubs, where technological innovations are widely accepted social media is quite a trendy platforms for all sort of activities. The number of posts every minute or even second ranges from reviews, shared pages, pictures, videos and comment on various issue and topics. All these information generated and growing bigger with time can be collected and analyzed to suit various business purposes. Using various social media strategies could lead to increase in sales, as it can be seen more than 50% of the participants surveyed, buy products seen on social media platforms with 25.71% engaging in such purchases very often. One of the trendiest movements is using consumers themselves for advertisement purposes. Campaigns, huge deals and sale offers attract a large audience who shares such posts with their friends on such social media platforms eventually leading to more sales figures and prices. As it is evident from the results produced above consumers are more comfortable and trust reviews posted by friends, family or other people within their network. The variety of languages in the region can also be seen on the data visualization charts; which calls for the requirement of multilingual analysis in the region. Posts in other languages cannot be ignored as they may containing vital information, reviews, feedbacks or comments on various aspects. Multilingual analysis is already a widely adopted analytical strategy, which is used by leading organizations and is a promising step to gain in-depth understanding.

4.1.2 Phase 2: Social media analytics and leveraging in organizations within the U.A.E

The second phase of the survey was to gain an understanding of how efficiently different sized organizations leveraged social media platforms, the techniques and reasons behind it.

**SMA being used within different sizes of organization**

Large sized organizations – Out of 35 participants from large scaled industries, 100% of the participants as shown in Figure 17 said their organizations were active on social media in the form of marketing, data gathering or analyzing. Being well established, these companies can be expected to understand social media marketing, its analytics and benefits in a deeper manner in comparison to medium-small industries.
Medium sized organizations – Figure 18 shows out of the 35 participants from various medium scaled industries 62.9% said their companies engage in social media marketing and its analytics to gain insights. Whereas, 37.1% of the participants responded their companies do not engage on any form of SMA.

Small sized organizations – From the 35 participants from small scaled industries as shown in Figure 19, 77.1% revealed that their organizations do not engage on any form of social media platforms to help improve business. Only 22.9% of the participants said that they do benefit from social media marketing and analytics. This was quite surprising especially within this region, where social media activity is quite a popular trend. The reasons behind such a large number not using social media within these small industries are explained in the following sections.

- Sectors where SMA is utilized in the region

Out of 105 participants, 61.9% said their companies engage on social media platforms and perform various types of analytics and research on data obtained from these sources. The pie chart in Figure 20 shows the various industries and their percentages of those engaging in Social Media and its analytics. As it is evident from the chart, the leading sectors from the sample population surveyed Technology, Media, Aviation, Banking, Health and Education sectors show active involvement on social media platforms. Sectors with small percentages are cumulatively put under the other category, which forms 7% of the chart.
Methods used by companies who use SMA

Out of 105 participants, 65 participants (cumulative number from large, medium and small scaled industries) said their companies engage in social media marketing as well as their analytics. These people were asked how their organizations or what were the methods their companies used to perform SMA. As seen in Figure 21 responses showed that, 40.63% of the participants whose companies engage in SMA said, it was outsourced or handed over to external parties. 31.88% said their organizations invested on in-house analytical software’s to perform such analysis and to gain insights. Whereas 37.5% said they did not know or were not sure on how their companies perform analytics on the data gathered through these social media platforms.

Effectiveness of SMA within organizations

Participants whose organizations performed social media marketing and its analytics were asked to rate how effective such techniques were for their organizations. As shown in Figure 22, 48% responded SMA is very effective for their organization, 41% said average effectiveness and 11% said it is not too effective. This varied response shows lack of proper strategy or could be even due to lack of planning or background research in the area.

Reason(s) provided by companies that do not use SMA

Out of the 105 participants, 40 survey participants (cumulative) from medium and small scaled industries said their companies do not invest in social media analytics. The participants were asked the reasons behind this and could choose multiple options. The number of various reasons selected by the participants were gathered and analyzed using a column chart shown in Figure 23. As per the chart, major reasons were concerns related to Return of Investment (ROI) and because the organization thinks SMA is not important for them.
These two major reasons were followed by future adoption plans. Other reasons included expensive, lack of knowledge and skilled labor which showed noticeable percentages regarding the reasons for companies not utilizing social media platforms.

Findings and discussions based on Phase 2 data gathering and analysis.

The second phase shows social media activity and analytics within the different sizes of organizations, out of which the large scaled industries utilize and benefit the most from it. Though medium scale industries show an active involvement, there are still quite a number along with small scaled companies who do not engaged in social media leveraging. There were various reasons behind this, out of which concerns related to ROI and the lack of importance given to SMA were the two major reasons highlighted in the data analysis. This indicates lack of awareness or knowledge in the field giving rise to such concerns. Organizations may not know what social media strategies to adopt in order to benefit their business hence causing worries about investing time and money in it. Marketing in today’s world does not happen without data and studying data, social media platforms and active data-growing platforms where the volume keep increasing with time. Proper guidance with context to social media marketing, analytics, data analysis, strategies could help these organization look at marketing in a different perspective and also increase their sales and profits. The larger percentage of the sample population surveyed responded that their organization involved in SMA and related activities. Technology, Media, Aviation, Banking and Health were the five major sectors identified to perform SMA in the region. It was noted that majority of the companies outsourced such activities and a very small percentage had in-house analytical software’s. The major reasons behind this could be lack of skill or lack of knowledge again. Some of
the participants were not even sure what techniques their companies used for SMA, revealing less employee engagement in such activities the importance of which was mentioned in Section. The other reason could also be a common stereotype that social media and related activities were concerns only of the Marketing department or of the CEO, and employees see no reason to involve in it. This common view among employees has been mentioned earlier in section 2.4.2 and is also proved from the responses obtained from the sample population in section 4.1.5. The effectiveness of social media in the organization showed a variance in responses where only less than half of the responses said SMA is very effective this could be due to reasons of improper planning, lack of knowledge how to leverage social networks properly or lack of skilled labor.

4.1.3 Phase 3: Participant’s understanding of social media analysis and its importance.

The last phase of the survey was to analyze the different levels of understanding of various participants with respect to SMA. In order to understand, a set of simple questions were asked as follows.

- **Are you aware that the information you enter on any social media platform/ websites can be collected and analyzed for various purposes?**

The question above helped realize the participant’s level of understanding about data that can be obtained from social media platforms. The responses are visualized using a pie-donut chart as shown in Figure 24. The feedback of responses showed 69.91% of the participants knew about data collection from social media platforms and its analytical use. 15.93% said they have a vague idea, which means they could have heard of it but is not too sure how or what its uses could be. 14.16% participants showed lack of knowledge with respect to this context.

- **Awareness regarding Social media analytics (SMA)**

In order to understand how aware the participants were about social media analysis three questions were asked where they could choose one option each. Out of 105 responses obtained, Figure 25 shows 42.86% which is less than half of the participants know what Social media analysis is. 38.1% are not aware about SMA and 19.05% just know or has heard about it. A cumulative percentage of participants who do not know or just has
heard about the benefits of SMA form 70.47% of the participant’s. Which is quite a huge figure compared to the mere 29.52% of those who know about its benefits. Furthermore, only 18.1% of the participants knows about various SMA analytical tools.

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>Only heard of it (or studied)</th>
<th>No</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you know what SMA is?</td>
<td>45</td>
<td>20</td>
<td>40</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>(42.86%)</td>
<td>(19.05%)</td>
<td>(38.1%)</td>
<td></td>
</tr>
<tr>
<td>Do you know about the benefits of SMA?</td>
<td>31</td>
<td>22</td>
<td>52</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>(29.52%)</td>
<td>(20.95%)</td>
<td>(49.52%)</td>
<td></td>
</tr>
<tr>
<td>Do you know or have heard about Social Media Analytical tools?</td>
<td>19</td>
<td>21</td>
<td>65</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>(18.1%)</td>
<td>(20%)</td>
<td>(61.9%)</td>
<td></td>
</tr>
</tbody>
</table>

These statistics obtained from these three question series shows a major lack of awareness among the participants even though most of their companies actively use social media platform for marketing, data gathering and analytical purposes.

**Importance of SMA in organizations**

Participants were asked how important or how beneficial would social media analysis be for their respective organizations. They were asked to choose from a scale of 1 to 3; where 1 stood for no (or minimal) benefits, 2 for few benefits and 3 extremely beneficial for the organization. After obtaining the responses, the answers were grouped according to the participant’s knowledge of SMA as per the previous section results. The groups were those who knew about SMA, who has minimal knowledge and who has no knowledge at all. The answers were then entered and visualized using a pie-chart representation and will be explained as below according to the three groups.

Group 1: Participants with knowledge about SMA

Out of the total 105 participants, 45 participants knew what social media analysis was. Figure 26 shows 91.11% of this group said SMA would be extremely beneficial for their organization. While 8.89% said it would have few benefits within their companies. This result is quite promising as employees knows the benefits SMA can offer with their varying level of understanding within the field. Naturally, these participants would also be part of social media activities within their organizations if they feel so strongly.
Group 2: Participants with less knowledge about SMA

Out of the total number of participants, 20 had only heard about SMA or has less knowledge comparatively. Figure 27 shows, 60% of this group said SMA is extremely beneficial, 30% said few benefits, while 10% said SMA has low or minimal benefits within their respective organization. This varying responses could be due to many reasons including insufficient knowledge, lack of active participation or involvement with social media strategies, less technical background and concerns related to investment or expenses. The various reasons or concerns related to the topic and those are common among the public are expressed in the literature review in section.

Group 3: Participants with no knowledge about SMA

40 participants out of 105 responded saying they did not know about SMA. In this group as shown in Figure 28, only 20% said SMA would be extremely beneficial for their organization. Whereas, the rest 57.5% said it would have a few benefits and 22.5% said SMA would have no or minimal benefit within their respective organizations.

Findings and discussions based on Phase 3 data gathering and analysis.

Analyzing participant’s level of understanding, about social media analysis was the main objective of phase three of this survey. Majority of the participants showed that they were aware that; data entered on social media platforms could be collected for analysis purposes. However, when asked if they had heard about social media analytics only less than half of the participants voted yes. Knowledge regarding the benefits of social media analysis or tools and techniques associated with SMA was noted to be poor among the participants. Lack of awareness or knowledge of SMA and its various strategies is quite evident in this study. This analysis supports the phase 2 findings, were participants showed major concerns related to ROI and thought that SMA would not be important for their organizations. The analysis of this phase showed differences in knowledge and therefore difference in opinions and various concerns related to SMA. It is natural for anyone to wonder if there would
be a relationship between the different knowledge levels related to SMA and the opinions related to them. In order to learn in-depth regarding this a hypothesis test was conducted and will be explained in the section below.

4.1.4 Hypothesis Testing

Based on the information collected through the survey, lack of knowledge or awareness about social media marketing and its analytics was clearly shown from the responses. It is also assumed from the data analysis that most of the concerns related to SMA regarding its ROI, expenses and importance could be a result of this lack of knowledge. A hypothesis testing was required to show the relationship between the participant’s view about SMA and its importance when compared to their knowledge level about the topic. The research problem was analyzed and tested using IBM SPSS, which is IBM’s power statistics and data analytics tool; that is also one of the market leaders for the last few years.

Research Problem: Participants with knowledge about SMA knows its benefits as well as it limitations and it is clear from their responses. Whereas those participant without knowledge seems to have a lot of concerns regarding the topic and shows a lack of interest or hesitance within the field. Does knowledge regarding SMA among people affect their understanding in realizing the importance of social media and its benefits within an organization?

Alternative and Null Hypothesis:

H₁ is the alternative hypothesis while H₀ is the null hypothesis for this research problem that will be subjected to Independent sample t-testing or two-tailed p test.

H₁ = There is a significant relationship between knowledge regarding SMA and understanding it’s importance or benefits within an organization.

H₀ = There is no relationship between knowledge regarding SMA and understanding it’s importance or benefits within an organization.

Alpha value = 0.05 (95% of the confidence level in SPSS). Which means the p value for the t-test obtained must be less that 0.05 in order to reject the null hypothesis, H₀.
Test Data:

The test data was obtained from phase three of the survey conducted. Out of 105 participants, 45 of them had knowledge about SMA, 20 had vague knowledge and 40 had no knowledge. For this hypothesis testing only 40 participant results from those with and without knowledge about SMA was taken in order to keep the same number of count in the both the groups being tested. The 20 participants data with minimal or less knowledge about SMA were not included since their level of understanding could not be measured accurately. Hence the groups to be tested are as follows and are also called as the grouping variables.

Group 1 – Participants with knowledge about SMA and it’s benefits. (40 no.’s)

Group 2 – Participants without knowledge about SMA and it’s benefits. (40 no.’s)

These groups will be tested on their responses on how beneficial they think SMA would be for their organization. The responses are obtained from the survey data and would be the testing variables. The mean from each group would be calculated and proceeded for hypothesis testing.

Test Results:

The screenshots of the test results obtained from IBM SPSS is depicted as in the Figure 29 and Figure 30 below.

![Figure 29- IBM SPSS Hypothesis Test Results showing count, mean and standard deviation.](image1)

![Figure 30- IBM SPSS Hypothesis Test Results showing 2 tailed p value and mean difference](image2)
Interpretations:

Figure 29 shows the number, mean and standard deviation between the groups with knowledge and without knowledge about SMA. The mean calculated shows the mean of the numbers of the responses obtained for benefits of SMA within their organization. Where 1 means least important, 2 stands for few benefits and 3 means maximum benefits. The mean has been created for both the groups, noticing the mean column (marked in green) a significance difference between the two means can be seen. This is could imply that the T test could reject the null hypothesis.

Further analysis is shown in Figure 30. Depending on the difference in variances the top or bottom column must be chosen for hypothesis testing purposes. IBM SPSS conducts a F test and the obtained p value is .004 (marked in yellow), which is less than the alpha value 0.05. Therefore, if .004 < .05, it can be stated that there is difference in the variance and hence equal variance cannot be assumed and hence the bottom column is selected for testing purposes (bottom column, marked in red).

While reading the bottom column results (marked in red), it can be seen that the t value is 8.054 and the p value of the t test is .000. If the p value is less than the alpha value then the null hypothesis can be rejected. In this case, the p value .000 < .05, hence the null hypothesis can be rejected.

Test Conclusion:

From the above independent t test, it is evident that there is a difference in the mean showing the benefits of SMA within the two groups tested. Hence, it can be concluded that there is significant evidence to support the research proposal. Therefore, there lies a relationship between participant’s knowledge regarding Social media analysis and his/her understanding about the importance of SMA and its benefits within an organization. This conclusion could also imply that those participants with knowledge on SMA may have concerns regarding data quality issues and measuring strategy success but would have no doubt on its prevalence to gain actionable insights. Whereas those participants without knowledge in the topic, naturally has concerns related to investments and shows reluctance to engage in this area.
4.1.5 Poll results

As part of the quantitative study two general polls was set and the same participants of the surveys were asked to take part in both the polls and the result visualizations are shown below.

- Poll 1 – Participants views about the various purposes of social media networks.

The participants were allowed to choose one or two options depending upon their views. The results shown in Figure 31 are depicted in a descending manner of options selected. Based on the poll responses, naturally most number of hits were to connect with family and friends which is ideally the basic need of social media platforms. But keeping in mind participants consisted of professionals with various backgrounds including Marketing, Sales and engineering; “all of the options” could have been a more suitable solution. Whereas, the hit count for the all of the above option is comparatively less showing ignore or lack of awareness about social media’s other benefits other than social networking. This ignorance could lead to the concerns related to ROI and wondering how social media activities could help benefit companies. In order to avoid this issue, companies can learn from leaders and other organization on how they effectively manipulate social networking platforms to gain valuable insights. Also companies must involve and engage employees in social media activities in order to educate them as well as obtain more innovative idea. Using proper strategies and conducting relevant back ground research will produce good results and sales figure with time.

![Purposes social media could be leveraged for?](image)

Figure 31- Participant’s views on various purposes of social media networks
Poll 2 – Participants views about social media responsibilities within an organization.

Participants of this poll were allowed to choose one but the most suitable option they think is majorly responsible for social media engagement within an organization. Although poll results as depicted in Figure 32, 32.38% of the participants felt strongly that all department were equally involved. Which looks promising as quite a good number does not believe that social media analytics and its responsibilities are solely only for the marketing team or of the Chief Marketing Officer (CMO). However, more than half the participant still believe in the prejudice or myth.

Employees need to understand that every department could utilize and benefit from the data or information obtained from these websites. In order to eradicate such a custom within professional environments, employees from various departments must be involved in social media strategies and planning. Teams could be built and involvement and participations in such campaigns could be promoted. When more number of people are involved, they directly or indirectly invite or communicate with more number of other people hence increasing brand activity as well as promotion. More engagement or activity means increased revenue and profits by the end of the year. All of these are connected to each other in one way or the other, a proper strategy aligned with company goals and planning is what is necessary within any organization before engaging in any form of social media analytics.
4.2 Qualitative data analysis – Focus Group Discussions

In order to pave way for deeper insights and understanding related to the SMA and the information obtained from quantitative analysis. A focus group discussion consisting of 8 participants was carried out over a period of two days. Participants for this discussion were selected on basis of their experience, knowledge and interest with reference to this topic. Participant’s roles and industry sectors are mentioned as per the table below. In order to keep up with the ethical and user privacy principles, personal information such as participant names, company names and similar data won’t be mentioned in the report.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Role</th>
<th>Industry</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Data Analyst</td>
<td>Technology</td>
</tr>
<tr>
<td>B</td>
<td>Business Development Manager</td>
<td>Technology</td>
</tr>
<tr>
<td>C</td>
<td>Marketing Manager</td>
<td>Airlines</td>
</tr>
<tr>
<td>D</td>
<td>Business Analyst</td>
<td>Airlines</td>
</tr>
<tr>
<td>E</td>
<td>Team Leader (Sales)</td>
<td>Retail</td>
</tr>
<tr>
<td>F</td>
<td>Social Media Manager</td>
<td>Retail</td>
</tr>
<tr>
<td>G</td>
<td>Marketing Executive</td>
<td>Food/Hospitality</td>
</tr>
<tr>
<td>H</td>
<td>Research Analyst</td>
<td>Education</td>
</tr>
</tbody>
</table>

*Table 3- Focus Group Discussion participants and their related designations and sectors*

4.2.1 Day 1: Focus Group Discussion regarding SMA and its uses and concerns in U.A.E

The first day discussion was to cover topics related to SMA benefits and concerns in the region. The 60 minute discussion covered 4 topics with a fixed time interval of 15 minutes each; where participants shared their views and the points were noted on sticky notes at the end of each session. These point will explained and summarized as follows.

- **General concerns related to social media analysis within the United Arab Emirates.**

All of the participants agreed that there is a significant gap within many organizations between social media information and obtaining business value from it. When compared to large organizations, this trend is more common between medium-small scale industries. Due to comparatively limited revenue within these organizations, they tend to stick to traditional means of marketing and business intelligence. Lack of awareness or poor knowledge within the organization on how to effectively leverage social media platforms is a common
limitation in these companies. Inability to communicate with the management about how social media platforms can bring profits and improved figures in the companies, rises concerns regarding ROI. All these concerns and issues may lead to think that social media analytics is complicated and is a big hassle and it is not important enough to invest in it. Participants came to these conclusions and opinions from previous experiences and dealing with clients and think organizations must be coached first on social media strategies and its data manipulation. Hiring skilled employees who have knowledge in this field is also an alternative. Organizations can also learn from the market leaders by studying their social media activities to understand how they attract and gain customers. Companies that are already benefiting from social media and its activities to could set up workshops and host seminars and act as role model in order to educate and increase knowledge on how to utilize and benefit from the digital revolution. On the other hand, companies engaging in social media analytics have constant concerns regarding the quality of the source and content of the data obtained for various platforms. Powerful tools have been invested upon to track source of information, for data cleaning and preparations in order to ensure the quality of the data before any form of analysis.

❖ Eradicating the stereotype “SMA is only for the Marketing team”.

Participants all agree that this is a general stereotype and common misunderstanding within the employees in any organization. But it was also agreed that this trend is changing within office environments in today’s world. Citing personal organizational activities, companies now involve one or more members from every department while planning social media strategies. This benefits both parties; as employees can be active participants and voice out their opinions, whereas companies gain more innovative ideas and learn new ways they can engage on these platforms. Because of various level and types of knowledge involving employees from various departments, a 360 degree view on the strategy regarding benefits, risks and analysis can be realized. When the social media strategy is in action, every department is given a responsibility or target to achieve, this could involve increasing campaign awareness, obtaining information about consumers, tracking campaign activity, building sales figures and charts. One of the participants even confided that their company keeps up employee participations by introducing competitions and similar activities. Such activities not only benefit the company financially, but also increases employee’s interactions as well as improves office environment. On the other hand, gain or loss from such activities on social media and its results, the responsibilities are shared equally within the company and is not burdened on just one department or person.
The need or significance of multilingual analysis within the region.

It was agreed that, any industry within the U.A.E is prone to have customers from U.A.E speaking in different languages as well as international customers. While setting up social media strategies, analyzing information in languages other than English cannot be ignored. Hence the requirement of multilingual analysis is a necessity not only in this region, but everywhere. A participant from the airline industry gave an example about their passengers commenting about their journey experience on social media platforms; apart from English a number of other languages were quite prominent. Hence analyzing tools chosen had multilingual analytical abilities in order to understand and realize sentiments and opinions related to passenger’s experience. It is a well-known fact that U.A.E is home to large number of expatriates hailing from different parts of the world and speaking various languages. It is also known that the native language in the region is not English but Arabic. Opening up any social media posts related to a brand posts in the form of comments or otherwise are seen in Arabic especially in this geographical location. Hence, it is important to take opinions in every language possible to be fair to all types of customers and to understand different needs and requirements. A strong positive opinion was voiced during the discussion for multilingual analysis within the region.

SMA benefiting the upcoming Dubai Expo 2020.

All participants first motive was to speak about U.A.E’s domination on social networks during the bidding stage back in the year 2013, where they won the bid hands down when compared to their competitors. The country understood the power of social media during their campaigns for the Expo bid, and now has over 1.3 million Facebook fans and 378,000 Twitter followers. With people from all over the world wanting to participate in the Expo, are already filling social media platforms with #UAE or #DubaiExpo2020. Participants feel that the organizers are already benefitting lot of information regarding opinions, ideas from posts or endorsements made on these social platforms and acts as a mode of communication with the rest of the world. Social media platforms has become an effective medium of obtaining quality content and to engage with potential visitors. Organizers already post on their official pages about upcoming attractions and activities related to the Expo, where they get bulk responses, comments and motivation which helps them learn and obtain new ideas. The participants also discussed about the major sectors of the country related to aviation, hospitality, tourism, government and others are also following a similar engagement on social media platforms in order to analyze and understand what the visitors and Expo participants expect of the country and its government.
4.2.2. Day 2: Focus Group Discussion regarding SMA limitations and visualization qualities.

In order to understand limitations that could affect visualizations obtained from social media analytics participants we shown two word cloud images. Word clouds are simple data visualizations that show the most common word associated with the keyword searched. The word clouds were produced using R Programming Language, which is a statistical programming language which is also common used to analyze data. The code used for all the three analysis is displayed in Appendix.

❖ Data mining using keyword #Expo2020 from Twitter

Participants were asked what they felt were the advantages and disadvantages for this visualization shown in Figure 33 containing 2000 results. According to the most frequently used terms the size keeps increasing. The positive reviews about this image was that it was colorful and easily understood by everyone and that it contains the most common terms related to the keyword. This will help analyzers understand topics associated to the keyword. As limitations, many were voiced out including a lot of terms don’t make sense such as “UUDUU”, inspite of data cleaning being performed a lot of such unnecessary terms affects the quality of the visualization. Hence the quality of data obtained was stressed to be the most important aspect and building block of any form of data analytics. Although words associated with Expo are shown, the relationships cannot be identified, for example it's difficult to find the relationship between the “jokes” for this particular keyword. Some participants even mentioned that since the data obtained was visualized on word cloud could be reason for limitations such as poor clarity, but quality of data still had to be verified whatever form of visualization was chosen.

❖ Sentiment analysis on keyword #Petrol from Twitter data

Sentiment analysis was conducted on Twitter data for the keyword petrol which is one of the major topic of interest especially within U.A.E. The sentiments were visualized as shown in Figure 34, according to their classifications on the word cloud. Participants thought that the image showed the most general sentiments or opinions associated with the keyword. For example the term “panic” associated with petrol was quite natural because the sudden hike in the price in the country caused panic among the citizens to an extent. The clear
classification of the sentiments helps viewers see words associated with various types of sentiments in a clear manner. But certain sentiments seems to be classified incorrectly but few participants believed that it could be the correct classification with respect to the whole sentence which we cannot see. In this visualization also lot of terms seems to make no sense to the keyword neither show how they are related to each other. A major drawback noticed was that certain sentiments were unclassified showing inefficiency within the analysis. But this is a general drawback of sentiment analysis and is still under research. Participant’s in general had a positive note towards sentiment analysis since they commonly agree that it helps in understanding a lot of customer sentiments and opinions with respect to various contexts. This helps companies reach out to those consumers in need of support or tend accordingly to their requirements in an effective manner.

- Limitations or issues related to SMA

From the discussion, majority of the participants believe that one of the major issues related to SMA is handling the bulk amount of data increasing with time. Due to the large amount of information generated over multiple social platforms companies have to invest huge amount in order to store such data, even the most powerful of the analytical tools cannot handle the issue of data volume yet. Another major issue discussed was that only a certain percentage of the consumers are active on social media platforms therefore only their opinions could be dominating on these social platforms. Leading to misconceptions from the analysis obtained, companies could miss out what the other half of its consumers are actually not happy with or require from the brand. Data quality is also a big issue in this context as quantitative and qualitative data obtained from these social media platforms needs to be measured. Social media data management needs to be practiced in order to produce high quality results. The limitations of sentiment analysis of not able to classify slangs, certain emotions also pose as limitations. Upon discussion, the group agreed that social media analytics have come a long way from when it became and being one of the trendiest topics of research could overcome all forms of limitations in the near future. A positive approach toward social media analytics and its future was noticed between the members of the discussion.
4.3 Framework for major Industry-Sector’s and their benefit’s using SMA

Different sectors depending on their types of customers, services provided and products have various requirements from social media actions and analytics. Sectors that have similar benefits from social media marketing and analytics are grouped below and the advantages from such social engagement and data gathering have also been mentioned accordingly. These sectors have been classified according to UK GVA and Employment by Sector, 2011 [68]. The classification table can be seen in Appendix F.

<table>
<thead>
<tr>
<th>Sectors</th>
<th>Social Media Marketing/Analytics Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge Services:</td>
<td>Knowledge sectors are more concerned about customer/user experience and how they can help improve their facilities through study of data obtained from various social media platforms, blogs, forums and online communities. Information obtained from these platforms helps build strategies to improve existing services as well as gain insights on customer expectations and concerns.</td>
</tr>
<tr>
<td>Education</td>
<td>• Get customer/student feedback about their experience with the company and its services through small surveys and polls on social networks.</td>
</tr>
<tr>
<td>Research &amp; Development</td>
<td>• After introducing new facilities or services user opinion could be collected in order to understand what has to be improved or removed and how happy the users are.</td>
</tr>
<tr>
<td>Communications</td>
<td>• Engaging in sentiment analysis could show how content and happy the users are with the organization and what they are expecting from the company. Accordingly actions can be taken and requirements could be satisfied improving the user’s experience leading to user satisfaction.</td>
</tr>
<tr>
<td>Business Services</td>
<td>• Have user-support on social platforms where users could convey their opinions, recommendations or issues with either a ‘#’ tag or ‘@’ sign which the company can view and obtain information. In this way the organization can respond and support users hence leading to customer retention and loyalty.</td>
</tr>
<tr>
<td>Banking/Financial Services</td>
<td>• Organize social campaigns with giveaways and prizes ensuring user activity and increasing brand awareness.</td>
</tr>
<tr>
<td>Media/ Digital, Creative &amp; Information Services</td>
<td>• Information collected of social media platforms could be used for research purposes in order to obtain insights about the current market trends and requirements. This could help in taking actionable decisions and strategies in order to increase business values as well customer experience.</td>
</tr>
</tbody>
</table>

Related studies and discussion:
Grunspan et al. research on using social network analysis to understand classrooms in order to inspire future educational studies by using the data collected [69]. Similar studies also show that adopting SMA can improve student experiences, reach out and stay connected with alumni as well as keep up the university spirit and brand image.
For research and development, a lot of researchers use data from various social media platforms through API’s connections to gather data to support their studies. Although lot of concerns or mixed views regarding the quality and integrity of the data obtained through those sources is present, many agree that the power of social media data is quite serious. An interesting research performed by the Oxford Internet Institute on the feasibility of using social data for research purposes shows its significance especially in obtaining public opinion on various topics [70]. Although the Beninger et al. of NatCen Social Research argues that, users those who voice their views publically should be consented and their anonymity should be confirmed while using their data for research purposes [71]. This problem arises since many users and their varied understanding of social media platforms and its uses may tend to post opinions not realizing that it could be open to research and other analytical purposes. The study urges and provides suggestions to researches to adopt other research practices to avoid such complications.

Social banking has become a common trend in order to focus and maintain customer-focused relationships. With the increase of social media utilization financial or banking services are expected to extend their services on similar platforms in order to improve customer experience, retention, improving brand reputation as well as for internal collaboration. Dr. Benedikter’s research on social banking and social finances shows the importance and significance of these techniques as a better system to handle money and finance [72]. Capgemini, a leading IT solution provider’s market studies implies that banks needs to develop a clear sense of understanding between key elements of social platforms, digital marketing and customer expectation in order reinforce customer engagement as well maintain loyalty [73]. Financial institution must not just settle with just creating a social presence but also leverage social media platforms in order to magnify customer information and to improve services.

Similar techniques and methods are being used in the other sectors under this classification to gain actionable knowledge from social media platforms and its leveraging.

Local Examples:

- Emirates NBD, one of the leading banks within the region engages on social media platforms to understand in depth about customer experiences as well as to solve issues faced by the users. The organization was named “The Social Brand of the Year” at the MENA Digital Awards 2014. Tools such as Google Analytics were also involved to enhance customer acquisition as well as to understand brand activity.

- Heriot Watt, Dubai campus engages on popular social media platform in order to increase student engagement in various activities within the campus and events organized by the university. The platforms also acts as a communication tool in order to understand the students opinions regarding various topics including their experience within the campus, faculties, staff and student groups.
These sectors involve the public has a whole and is directly or indirectly affecting or benefitting a country. Therefore, the public’s opinions and engagement is essential for these sectors and the most effective way is to communicate using social media platforms.

- Public’s opinions and sentiments over various topics can be understood and analyzed over social media platforms, blogs and forums.
- While introducing new attractions, events and public utilities, the public can be included during the planning and development stages in order to obtain new ideas as well as to keep the public content.
- Social media engagement could improve communication between the Government and the public and their collaboration. It also leads active participation of the public and adds transparency between public and the state activities.
- Studies show that sentiment analysis can help police as well as cyber security to analyze groups and individuals associated or instantiating illegal activities. Hence SMA even helps in protecting the public.
- Political campaigns and bids on social platforms shows the public’s and the world’s opinion and helps with election events.

Related studies and discussion:

The impact of social media on the public sectors can been seen in Karakiza’s recent research on how effectively government bodies leverage on social media platforms [74]. The study also proposes an open government model that can be adopted to obtain maximum benefits for the government sectors. Examples of best practices within the context have been sited from all over the world. However, the research is concluded saying that the social media applications and its adoption is at an early stage within government sectors. This is because they are more focused on public participation and transparency and less on collaboration techniques. Murphy et al. study further justifies leveraging social media platform by government sectors to gain knowledge on public opinion [75]. With the increased use of technology, popularity explosion of social media and then trend of publically posting opinions make social networks the most preferable medium to obtain information. The study further looks into data quality validation, legal and ethical considerations as well as well addressing sampling challenges. Proposals on improved integration between surveys and social platforms, proper manipulation of social media and its unique features have also been looked into. Picasa-Vela et al. study agrees with both the above research, which means engaging on social media platforms for governmental insights has its benefits as well as drawbacks [76]. The public would be delighted with the transparency offered by the government in their activities and encouragement to involve in order to improve the face of the government. Whereas, this could lead to a risk of mistrust also when citizens learn that the government is prying on each content individually which
could pose as a threat for certain individuals. Although the study agrees that such involvement with the public could change the culture of communication with the citizens, however also points out at the fact that at some point, control over the quality of the information could get lost. The study suggests that public sectors must have strong social media strategies with a solid plan in order to leverage efficiently while handling the risks associated with it.

Studies also show that analysis of data obtained through social networks could help identify crime patterns, predict harmful events as well as help fight crimes. Bolla’s research within the topic could relate dangerous data obtained from twitters feeds and connect them to crime scenes [77]. Thereby enabling prediction of crime patterns within a particular region as well predict the probability of similar occurrences. Sentiment analysis was performed on the Twitter metadata and based on the sentiments posts were classified as either safe or dangerous, over 100,000 crime related tweets were collected using this technique. The research also depicts that with problem social network monitoring real time crime patterns could be depicted in order to prevent crimes even before they occur. Similar studies of Wang et al. [78] and Corso et al. [79] also agree and suggest that social media analysis could lead to predictive crime analysis, show hotspots as well as various risk levels. Apart from social networks, various online platforms and websites on hidden IP addresses can also be tracked for illegal activities. Such online medium which harnesses illegal activities and trade are collectively called as the Dark Web. It is a collection of a large number of networks which resorts to various illegal techniques to hide their IP addresses or identities. A normal user can view such networks but the identity of the person or the team behind such network remain unidentified. These markets on the other side of the internet are flourishing from their profits from illegal trade of drugs, child slavery, pornography and similar crimes. Proper monitoring and analysis of such online platforms could provide clues, connections and evidences to track the culprits or criminals.

Political communication and engagement has become quite prominent on social media platforms in the recent years. Political campaigns have been hosted encouraging participation in order to win votes from the public for the best candidates. Stieglitz and Dan-Xuan’s study on social media with respect to it turning into a political channel provides a wide aspect in this area [80]. It shows that this helps political organizations communicate with increased transparency with the public as well as helps with political decision making processes. The research points out that due to lack of knowledge among politicians about proper leveraging on social media platforms the benefits obtained is not at full potential. The paper also suggest a strategy proposal for political communication consisting of data tracking, analysis and corresponding analysis methods in order to obtain deep insights and to take practical decisions.
Local Examples:

- U.A.E in its successful bid in hosting the Expo 2020, realized the power of social media and its importance for generating interest during the campaign. Even after winning the bid, the country continues to engage on all major social platforms to include the public in its planning, activities and upcoming attractions. Social networks serve as a platform to obtain new ideas and opinions from the public as well as to reach participants from all over the world, being an international event. The organizers also hold a number of social media campaigns and contests in order to ensure the public’s engagement.

- The Dubai police is known for monitoring social media platforms round the clock and states that monitoring of these platforms are not considered as violation of public’s freedom [81]. The information obtained from various platforms are retained in the server for 18 months and is said to help the police to track the source of various illegal activities.

Other Services:

- Hotels & Restaurants
- Technology
- Transport, Storage & Distribution
- Fashion/ Retail
- Health & Social Care
- Community, Social & Personal Services
- Real Estate
- Support Services

These sectors along with customers experience are also concerned about user demands for various specifications, experiences and designs.

- Increase brand awareness by increasing the number of fans and followers on various social platforms in order to increase sales.

- Depending on opinion mining and text analysis could provide insights on customers’ requirements. Thereby helping production planning and supply chain activities.

- Media, sports, fashion events and other campaigns can increase number of audience and participants by engaging on social media advertisements which show fun activities, prizes and interesting events. Also, engaging the audience while planning such events can provide innovative ideas by encouraging them to share what they think would be exciting and interesting. This also provides as a good communication and engagement strategy.

- Video’s and pictures of events can be posted on social platforms especially Instagram, Facebook and Twitter in order to show followers and fans a glimpse of the fun activities and how interesting such events can be, hence increasing audience in the upcoming events.

- Gather socially active audience by making them take part in online competitions, e-commerce shopping, online sale and similar activities will lead to user interaction, improve sales as well gain loyalty to some extent.

- Social networks are also an effective source of information on how competitors are manipulating their audiences and the different tricks and trends of social media marketing.

Related studies and discussion:

In health and life sciences sector social media data and its analytics provide information on common diseases and understanding it’s behaviors and patterns among the people and needs for various types of medicines and their demand. Apart from this, social media engagement also helps to share information about certain diseases and their awareness with
the public, it also acts as a platform of collaboration and innovation - between medical science and IT for example, feedbacks and responses on patient experience in various areas. IGATE’s research paper [82] on the topic agrees with the above statements and believes that though measurement of ROI is difficult, visible benefits can be seen from social networks and its leveraging.

The food industry benefits from the social media in various ways such as learning about customer experiences, their tastes as well as provides as an effective platform of communication. Rutsaert et al. study shows that social media also acts as a statistic tool in order to communicate about health benefits and risks related to food [83]. A number of Facebook pages and Instagram profiles can be seen promoting healthy food and eating which falls under this form of communication. Case studies and research conducted in the pizza industry by He et al. also shows that social networks are not just for promoting the brand or for advertisement purposes but can also provide information to make strategic decisions [84]. The studies shows that social media platforms also provides as a positive communicating channel with customers as well as maintain bonds.

The airline industry has been active and enjoying benefits from quite a number of years engaging in sentiment analysis, tracking brand awareness, understanding customer experience and track competitors social media engagement. Gaining new customers through these social network have also been one of the airline industries objectives by providing attractive offers on their official pages as well keeping existing customers happy and satisfied. Leung et al. book explains in detail how airlines could leverage social network platforms in order to attract passengers by active user engagement and user involvement [85]. Other sectors of hospitality, tourism and banking engage in similar strategies and social media techniques to increase their business values as well as improve their customer satisfaction.

The fashion industry is definitely a sector enjoying the maximum benefits of social media marketing since millennial generation are their targeted customers. Fashion lovers including celebrities and public figures enjoy flaunting their new outfits on various social media platform, influencing their fans, followers, peers or friends to follow the trends. Social networks also act as an efficient communication medium in order to understand the trends in demand which can further help in production planning. Kim and Ko’s empirical study on social media marketing and its effects on luxury fashion brands shows that proper engagement techniques on social platforms can lead to increased customer equity [86]. The importance of understanding and managing customers on social platforms in order to enhance retention and loyalty due to increased competition in the market is also emphasized in this study.

Similar practices could be adopted by various sectors in order to increase brand activity and image, understanding consumer and purchasing behavior as well the current trends and requirements. Thereby earning customers who are the most valuable assets for any organization in order to achieve increased business values.
Local Examples:

- Emirates Airlines, one of the world’s leading airlines owned by the Dubai government engages actively with their customers on social media platforms in various manner. With over 4.7 million fans on Facebook and 5 million followers on twitter follows brilliant social media strategy which has lead them to strive and become one of the world leaders in their sector.

- Bosch Middle East, one of the branches of the leading multinational German company also engages on social media platforms to attract potential employees, partners, supplier and customers. They are actively involved in Facebook, Twitter and Instagram where images and videos of their new products, services or events are uploaded. Also reviews and issues on various social platforms are continuously monitored to provide prompt customer support in order to obtain their loyalty.

- Aster DM Healthcare, is one of U.A.E’s leading health care services. Their engagement on social media includes spreading awareness about various diseases among their fans and followers, providing useful tips as well as promote good living. Apart from these, user engagement is employed by a number of competitions, offers, campaigns and events on various social media platforms as well as on their own official website. Reviews and comments about patient experience is also encouraged on various social platforms and forums.

Table 4- UK Classification of sectors and related SMA benefits
Chapter 5 – Evaluation

This chapter is a reflection of the entire research and the report as a whole where the different chapters and methods used for this study are evaluated and explained accordingly.

The first method of research was the study and critical evaluation of relevant literature in accordance to SMA. The literature review resources mainly consisting of books, journals, articles and online documents were from well validated sources hence ensuring the quality of the research material. Source of research also includes databases of Emerald, Science Direct and IEEE. All resources used has been cited and referenced respectively. The literature review discusses major topics relevant to this study such as the influence of social networks, the various types of analysis and its visualization, the need for multilingual analysis and the requirement of SMA in the world. Not only benefits related to SMA have been elaborated on but also related concerns and limitations hence showing the study was well biased. Examples of leaders from various sector have also been provided so that the reader gets a clear perspective of the topic and can related them to real life situations.

For the case study purpose, surveys and focus group discussion were conducted in an orderly manner keeping in mind user data protection, identity protection and participant’s time as well ethical agreements. Survey participants consisted of mid-level management as well as junior employees, as the study also revolved around employee’s perspective or awareness about social networking and its analysis. The quality of the data obtained from these survey cannot be measured in reality, but is assumed that all participants have participated in a sincere manner. Focus group discussion consisted of experts in the field of SMA and data analysis in order to understand topics related to research. This discussion also worked as a support to back up the quality of the research results and findings.

The data obtained during the quantitative as well as qualitative research were stored with respect to data protection laws in order to preserve the participant’s identity. Data was encrypted and stored during the analysis phase with password protection so as to avoid any legal troubles in future. The data collected will be discarded once the project phase comes to an end. Consent forms have signed by all the participants and have been assured that the user’s anonymity will be maintained throughout. During discussions, personal examples were mentioned and used to make clarity of the topics. Thereby it was ensured that no company names will be mentioned anywhere in the report in order to adhere to contractual laws.

The data obtained has been analyzed and visualization in very simple manner in order to ensure user readability and understanding. The results were clearly explained with tables and figure with clear textual discussion. All results obtained are with respect to the sample population and it must be kept in mind the population surveyed is only within the U.AE and discussions concluded could vary according to time and location. The research
hypothesis has been derived from the studies and according to the data collected from the survey it has been stated, tested and proved.

R programming was only used as a tool and to obtain a better understanding of social data mining. The code used for creating the images was obtained from RBloggers and ThinkToStart, these have been referenced. The code was not build buy the researcher as it was out of the scope of this project as well as for the researcher’s understanding. However, studying various codes has made the learning process very constructive. Although the R-code consisted of data cleaning techniques before the visualizations were created, it still could not get rid of ambiguous data as has been mentioned in the limitation in the above sections. The twitter data used for the analysis were all public posts and were obtained via the public API’s. Since most of the code available on the blog were for older versions of R applications, a lot of changes had to be made to the code after a certain amount of research.

The SMA-benefit framework consisting of various industries does not include the manufacturing sectors since the study was conducted within the U.A.E. Manufacturing industries are not present in a large scale within the region especially within the sectors of technology, automobiles, fashion and medicine. Local examples as well as related studies have been cited so as to support the justifications and for the reader’s better understanding.

The quality of the report has already been evaluated using peer reviews. Three readers from various backgrounds were made to read and comment on the report in order to make sure the report was not too technical but kept the academic standards. Images within the report are kept small but legible keeping in mind the number of pages. Most examples given or made throughout the report are using local examples since the case study was conducted within the U.A.E. The study is in general targeting researchers, professional or general reader who want to learn and understand about SMA and its basics.

5.1 Limitations and Future Work

The sample population for the survey was restricted to 105 participants, with 35 participants each from large, medium and small scaled industries. This was because the targeted participants were selected based on their willingness to participate as well as it had to be ensured that they participate sincerely in order to maintain the quality of the data obtained. The limitation of time was also a major reason, according to the research plan not more than 3 weeks could be allocated for survey data collection. The studies and the participants are limited only within the U.A.E and therefore the results obtained for this research could vary with respect different sample population in other geographical locations.
The demonstration of multilingual analysis using R programming could not be achieved due to lack of lexical resources and no advanced knowledge in the programming language. However, related studies especially with respect to mining Arabic texts have been mentioned in the literature review.

Further work in this field can be proposed as following:

- A detailed study of multilingual analysis including a demonstration of analyzing sentiments involved on a particular topic in three or more languages. This will help determine the variations and similarities of opinions in English and other languages. The research will either be able to point out similarities and difference and how these variations can make a difference.

- Conduct different forms of SMA including sentiment analysis and context analysis using various applications and analytical tools to create a framework and to understand the advantages and disadvantages over one and other. This study could comparisons as well as advantages and similar disadvantages that are common among these tools and propose methods overcomes these limitations.

- Conduct case studies on companies using SMA and those companies using traditional methods of marketing and communication. Advantages and disadvantages on both sides could be pointed out which could maybe lead to collaborations proposals in order to obtain much effective techniques.
Chapter 6 – Conclusion

Social media engagement among the public was well clarified, after this study showed how influential social networks are within the United Arab Emirates. Being a country where latest technologies are widely adopted and thriving, the use of smartphones make the influence of social media much prominent within the region. The variety of languages was also seen, especially within this region leading to the conclusion that multilingual analysis is more of a necessity than a requirement in such a geographical location consisting of large cultural diversity. From the literature reviews and discussion based on the topic the importance of analyzing sentiments as well as opinions in languages other than English can be realized. It can hence be stated that all forms of organizations that have customers speaking in different languages or are from different places must invest in multilingual analysis in order to understand various requirements, opinions and trends among the people and could also prove a provide a much clearer perspective of things.

Research in the field also showed a large purchasing power over social networks within the region. Despite of this fact almost 40% of the participants surveyed responded that their organizations do not involve in social media analytics. An insight into the topic showed that most of such negative responses were from medium to small scaled industries having less than 50 to 200 employees. The reasons behind not adopting the technique majorly included concerns related to ROI, lack of knowledge and thinking that this method has no importance for their organization. These reasons show ignorance and lack of awareness within the topic which could lead to conclusions and fears related to those mentioned above. Hypothesis test conduct for this study proved that the lack of awareness among the people could lead to such concerns and a notion that such techniques will not be effective for the organizations. The other 60% of the participants said that their companies leverage social networks, utilize information obtained from it, makes decisions accordingly as well as reach out to customers in a more efficient manner. However only 47% from these said that social media analytics has been very effective in their companies turn overs, customer support, increasing audiences and for increasing profits. This factor highlights comparatively poor strategies for social media marketing and its analysis within these organization. As already elaborated in the literature review in section 2.7 every action on social platforms such as tweets, likes, comments, fans and followers should all be guided by a comprehensive plan and must be part of a strategy in order to achieve predetermined goals. Strong and effective strategies must be planned in advanced including decision makers from various department of the organization. This will not only lead to innovative ideas but will also help in eradicating the myth that social media and its analytics within an organization is solely the responsibility of the marketing team or the CMO. When all the employees take an efforts and participate in such a strategy it make it more efficient as social networking is about people and communication. Therefore more people or customers are attracted when more people or employees are involved and take an initiative to support the movement.
Various analysis and tools that can be used for social network analysis as mentioned throughout the report are available in the market. It is important for organization to first and foremost figure out what they expect to gain or achieve from SMA and then choose tools and techniques accordingly. Different sectors of industries could have different types of needs and requirements from social networks as seen in Section 4.3. Hence, it is essential to perform a good background research and study within the area by researching on leaders and competitors. This could not only show different types of methods but also show those in trend and create valuable results. Social media is one of best and an effective source of communication between organizations and their consumers. Also important information regarding sentiments and opinions associated to the brands or its products and services can also be determined. This knowledge could help with decision making and planning business operations in such a way in order improve customer satisfaction as well as reaching out to those in need of support. Engagement on social networks could also be a self-reflecting experience which could show the brand popularity as well as its influence on the people.

Although there are limitations and concerns related to social media and its analytics it must be remembered that there is a drawback in every form of technology and its related activities. Even traditional forms of business activities and marketing methodologies have their limitations and associated risks. Therefore, it is important to thoroughly understand the field and then find way to counter such issues or concerns in an effective manner. Finally, the research concludes that social media analytics is one of the powerful methods of reaching out to people as well as understanding their thoughts, actions and requirements in today’s world. An increase in awareness or interest to venture into the field could show hidden insights and values to any form or sector of industries.
References


Bibliography


Appendix

Appendix A – Survey Form

- The information page which gives the participant’s an overview of the survey and its purpose.

Hello,

You are being asked to take part in a survey that would provide necessary information that is required for a Post-graduate Research Study. Before you begin, kindly read the consecutive information. Depending on what you read, you may decide if you wish to participate or not.

The purpose of this research is to understand how many professionals or individuals understand the purpose and need of Social Media Analytics (SMA). It is fact that majority of the population especially within the United Arab Emirates are active on various social media platforms and use it for different purposes and this survey could help prove that. Organizations use it for marketing purposes, advertising, target potential customers and similar other purposes. But how many individuals/organizations actually invest or even try to understand Social Media Analytics? Being in a geographical location such as the UAE with multi-cultural diversity, the importance of multilingual analysis will also be studied. Also understanding purchase power on social networks and various levels of understanding of SMA among different participants will also form the objectives of this survey.

Your confidential answers will help me understand various topics with respect to the topic as well as define a research hypothesis in context to this topic. Please be noted that no personal information such as name or company name is asked anywhere within the survey. Although, the sectors under which your organization falls under is asked in the second phase since it is critical for this research. The survey is simple and will take maximum of 20 minutes from your schedule, hence please go through all the questions and options as your sincerity could make a big difference for this research.

The survey is divided into three phases:

**Phase 1 - Influence of Social media**
This section will help realize how active people are on social media and what kind of posts do you see often. The various languages you see on social media. For example, Arabic posts are very common within the Middle East. Also how engaged you are on social media platforms and what kind of posts convince you the most to indulge in such activities.

**Phase 2 - Marketing on Social Platforms**
The sections helps realise depending on the size of the company, how many organizations of various sectors, market or advertise on such platforms. Do they perform any analysis on social networks to understand better about the customer or products? What kind of tools do they use if they are already performing such analysis and how effective are such techniques for the organization. If they do not, why and what are the reasons behind it.

**Phase 3 - Understanding Social Media Analysis**
This phase will understand the various level of understanding between different participants on social media analysis and how it can benefit organizations. Their opinion about the importance of SMA in an organization will also be understood in this phase.

In case of any queries please do not hesitate to contact on the email address provided below. A BIG thank you in advance.

Kind regards,
Chantelle Stephen
MSc Computer Systems Management
Heriot Watt, Dubai Campus
Email: cx373@gwu.ac.uk
The consent form for the participant before proceeding to the survey

Consent Form

Subject: An Insight into Social Media Analysis
Name of Researcher: Ms. Chantelle Stephan
Name of Supervisor: Dr. Hind Zantout
Student ID: H00200852
University: Heriot Watt, Dubai Campus

I have understood all the aforementioned information provided by the researcher and have consented to participate in the survey.

I understand that my participation is entirely voluntary, the data collected during the research will not be identifiable, and that I have the right to withdraw from the project at any time, without any obligation to explain my reasons for withdrawing.

I further understand that the data I provide may be used for analysis and subsequent publication, and provide my consent that this might occur. I have been given contact details of the researcher in the information sheet.

By clicking the Next Page button, I agree to all the information above.
Phase one of the survey to help understand the influence of social media within the U.A.E

### Phase 1 - Influence of Social Media

1. How often do you use/check Social Media? (Social Media involves all platforms such as Facebook, Twitter, Instagram, Pinterest etc. as well as micro blogs, vlogs, YouTube-ing etc.)
   - Every few minutes
   - Daily
   - Never
   - Every Hour
   - Weekly
   - Monthly

2. What kind of posts do you usually see/read on social media platforms?
   - Advertisements
   - Pictures
   - Videos
   - Friends reviews/posts
   - Shared pages

3. Other than posts in English, what other languages do you see/post on your social websites?
   - Arabic
   - European
   - Indian Languages
   - Other

4. How often do you purchase products or services that you see on Social Media Platforms? (Online purchasing or purchasing same products from department stores.)
   - Never
   - Sometimes
   - Often
   - Very Often

5. What form of posts would convince you to buy or invest in a product or service of interest? (Choose ONE most convincing post from the options, choose others if option not mentioned.)
   - Peer Review/Professional Reviews
   - Brand Advertisements
   - Attractive Sale Offers
   - Other
Phase two of the survey in order to understand the engagement of social media analytics within the region

Phase 2 - Marketing/Advertising on Social Platforms

6. Which size-category does your organization fall under?
   - Small scaled (>50 employees)
   - Medium scaled (50-200 employees)
   - Large scaled (<200 employees)

7. What sector(s) does your organization fall under?
   - Technology/Engineering
   - Construction/Civil
   - Hotel and Hospitality
   - Other
   - Education
   - Fashion/Retail
   - Banking/Financial Organizations
   - Government/Related Organization
   - Health/Related fields
   - Aviation/Transport Related
   - Sports/Related fields
   - Media/Entertainment and Events

8. Does your organization advertise or carry out any other form of marketing strategies via social media platforms?
   - Yes
   - No
   - I don’t know

9. If "YES" to Question number 8: How does the organization perform such analysis?
   - In-house software
   - External parties
   - Not sure what techniques are used

10. If "YES" to Question number 8: On a scale of 1-3 how effective is social marketing and it’s analysis for your organization?
    (1 = Not too effective, 2 = Average, 3 = Very Effective)
    - 1
    - 2
    - 3

11. If "NO" to Question number 8: Why do you think Social Media Analysis is not conducted within your organization?
    - Lack of knowledge
    - Concerns regarding Return of Investment (ROI)
    - Lack of skilled labor/ expertise
    - I don’t think it is important
    - Thinking of adopting such techniques in the future
    - Too expensive
• Phase 3 of the survey in order to understanding the awareness of social media analysis among the participants. This is the last phase followed by the “Thank You” page.
Appendix B – Polls

- Poll created in order to find participants views on how social media platforms could be leveraged.

- Poll created in order to find participant’s view on who they think Social media engagement is responsible within an organization.
Appendix C – IBM SPSS (Hypothesis Testing)

- Data in IBM SPSS ready for analysis and testing.

- Independent samples T Test highlighted in red is selected for two tailed testing
Step where the test variables are chosen, in this case it would be the rating given by the participants on a scale of 1 to 3 regarding the importance of social media marketing and its analysis within an organization. The group variable are the two groups of participants, Group 1 with knowledge and Group 2 without knowledge about SMA.

Defining the groups for testing where Yes = Group 1 = with knowledge of SMA and No = Group 2 = No knowledge about SMA.
The confidence interval percentage is chose as 95% which means that the alpha value is 0.05 (standard value).

T-Test results obtained after hypothesis testing.
Appendix D – Focus Group Discussion

- Consent form provided to selected participants before the focus group discussion. This document also includes information regarding the purpose, topics and duration of the meeting.

---

**FOCUS GROUP DISCUSSION**

Dear Participant,

As discussed earlier, you have been selected to be part of a Focus Group Discussion which will be held over a period of 2 days. The conclusions and findings obtained during this discussion will be used for Qualitative analysis of a post-graduate research on Social Media Analytics (SMA). The discussion days have been chartered as follows:

**Day 1:** SMA and its uses and related concerns. (60 minutes)
- General concerns related to social media analysis within the United Arab Emirates.
- Eradicating the stereotype “SMA is only for the Marketing team”.
- The necessity of multi-lingual analysis within the region.
- SMA benefiting the upcoming Dubai Expo 2020.

**Day 2:** SMA visualizations and related limitations. (30 minutes)
- Evaluating 2 word clouds based on related words and sentiments respectively.
- Limitations of SMA.

You have been selected based on your experience and knowledge within the areas of marketing, data analytics and social media analytics. Your views and opinions on these topics will help with the research project and its objectives. Please be noted that no personal information related to participant name, organization number or any kind of identity will be revealed or used within the project or its report.

“I agree that my participation is entirely voluntarily, the information obtained during the discussion will not be identifiable and I have the rights to rise a legal complaint if any contractual laws have been violated by the researcher.

I further understand that the information or findings obtained during this discussion will be used for studying and subsequent publication and provide my consent that this might occur. Contact details of the researcher has been provided.”

By signing below, I agree to the information above.

Sign: _______________________________
Appendix E – R Programming

- The code obtained from RBloggers is updated with a few changes in order to be compatible with RStudio v3.0.3. Twitter OAuth was created and verified in order to obtain the Twitter data over the public API's. The code was then run step by step in order to produce the world cloud. The world cloud used in the report is different from what is seen on the R Console since it was created on an earlier day and an enlarged image has been shown below.

- Word cloud created using different words associated with #Expo2020
- **R Programming code to produce word cloud with Twitter data for words associated with keyword #Expo2020 [63]**

```r
library(twitteR)

tweets = searchTwitter("#Expo2020", n=2000)

library(plyr)

tweets.text = laply(tweets,function(t)t$getText())

clean.text <- function(some_txt)
{
  some_txt = gsub("&amp","", some_txt)
  some_txt = gsub("(RT|via)((?:b\W*\@\w+)+)", "", some_txt)
  some_txt = gsub("@\w+","", some_txt)
  some_txt = gsub("[:]punct:]", "", some_txt)
  some_txt = gsub("[:]digit:]", "", some_txt)
  some_txt = gsub("http\w+", "", some_txt)
  some_txt = gsub("[ t]{2,}"", "", some_txt)
  some_txt = gsub("^\s+||\s+$", "", some_txt)

  # define "tolower error handling" function
  try.tolower = function(x)
  {
    y = NA
    try_error = tryCatch(tolower(x), error=function(e) e)
    if (!inherits(try_error, "error"))
```

84
y = tolower(x)

return(y)

}

some_txt = sapply(some_txt, trytolower)

some_txt = some_txt[some_txt != ""]

names(some_txt) = NULL

return(some_txt)

}

clean_text = clean.text(tweets.text)

library(tm)

tweet_corpus = Corpus(VectorSource(clean_text))

tdm = TermDocumentMatrix(tweet_corpus, control = list(removePunctuation = TRUE, stopwords = c("machine", "learning", stopwords("english")), removeNumbers = TRUE, tolower = TRUE))

library(wordcloud)

require(plyr)

m = as.matrix(tdm) #we define tdm as matrix

word_freqs = sort(rowSums(m), decreasing=TRUE) #now we get the word orders in decreasing order

dm = data.frame(word=names(word_freqs), freq=word_freqs) #we create our data set

wordcloud(dm$word, dm$freq, random.order=FALSE, colors=brewer.pal(8, "Dark2")) #and we visualize our data
The code obtained from RBloggers is used updated with a few changes in order to be compatible with RStudio v3.0.3. Twitter OAuth was created and verified in order to obtain the Twitter data over the public API’s. The code was then run step by step in order to produce the world cloud. The world cloud used in the report is different from what is seen on the R Console since it was created on an earlier day and an enlarged image has been shown below.

- Wordcloud showing sentiments related to keyword #Petrol
• R Programming code to produce word cloud with Twitter data containing sentiments obtained for keyword #Petrol [64]

# required packages
library(twitteR)
library(sentiment)
library(plyr)
library(ggplot2)
library(wordcloud)
library(RColorBrewer)

# harvest some tweets
some_tweets = searchTwitter("#Petrol", n=3000, lang="en")
# get the text
some_txt = sapply(some_tweets, function(x) x$getText())

# remove retweet entities
some_txt = gsub("(RT|via)((?:\b\W*@\w+)\")", "", some_txt)
# remove at people
some_txt = gsub("@\w+", "", some_txt)
# remove punctuation
some_txt = gsub("[[:punct:]]", "", some_txt)
# remove numbers
some_txt = gsub("[[:digit:]]", "", some_txt)
# remove html links
some_txt = gsub("http\w+", "", some_txt)
# remove unnecessary spaces
some_txt = gsub("\s\{2,\}", "", some_txt)
some_txt = gsub("\s+\s+$", "", some_txt)
# define "tolower error handling" function
try.error = function(x)
{
    # create missing value
    y = NA
    # tryCatch error
    try_error = tryCatch(tolower(x), error=function(e) e)
    # if not an error
    if (!inherits(try_error, "error"))
        y = tolower(x)
    # result
    return(y)
}

# lower case using try.error with sapply
some_txt = sapply(some_txt, try.error)
# remove NAs in some_txt
some_txt = some_txt[!is.na(some_txt)]
names(some_txt) = NULL

# classify emotion
class_emo = classify_emotion(some_txt, algorithm="bayes", prior=1.0)
# get emotion best fit
emotion = class_emo[,7]
# substitute NA's by "unknown"
emotion[is.na(emotion)] = "unknown"

# classify polarity
class_pol = classify_polarity(some_txt, algorithm="bayes"
# get polarity best fit
polarity = class_pol[,4]

# data frame with results
sent_df = data.frame(text=some_txt, emotion=emotion,
                      polarity=polarity, stringsAsFactors=FALSE)

# sort data frame
sent_df = within(sent_df,
                 emotion <- factor(emotion, levels=names(sort(table(emotion), decreasing=TRUE))))

# separating text by emotion
emos = levels(factor(sent_df$emotion))
nemo = length(emos)
emo.docs = rep("", nemo)
for (i in 1:nemo)
{
    tmp = some_txt[emotion == emos[i]]
    emo.docs[i] = paste(tmp, collapse=" ")
}

# remove stopwords
emo.docs = removeWords(emo.docs, stopwords("english"))

# create corpus
corpus = Corpus(VectorSource(emo.docs))
tdm = TermDocumentMatrix(corpus)
tdm = as.matrix(tdm)
colnames(tdm) = emos

# comparison word cloud
comparison.cloud(tdm, colors = brewer.pal(nemo, "Dark2"),
                 scale = c(3,.5), random.order = FALSE, title.size = 1.5)
Appendix F - Images

- Enlarged image showing LinkedIn connections and nodes as mentioned in section 2.2.3
- UK Classification of Sectors according to BIS Analysis (2011).

<table>
<thead>
<tr>
<th>Low-Med Tech Manuf.</th>
<th>Output (GVA)</th>
<th>Employment</th>
<th>% Share</th>
<th>% Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food, Beverages &amp; Tobacco</td>
<td>27,771</td>
<td>2.0%</td>
<td>399</td>
<td>1.3%</td>
</tr>
<tr>
<td>Metal, plastic and non-metal mineral products</td>
<td>28,005</td>
<td>2.0%</td>
<td>584</td>
<td>1.9%</td>
</tr>
<tr>
<td>Other Manufacturing</td>
<td>21,040</td>
<td>1.5%</td>
<td>505</td>
<td>1.8%</td>
</tr>
<tr>
<td>Shipbuilding</td>
<td>1,246</td>
<td>0.1%</td>
<td>32</td>
<td>0.1%</td>
</tr>
<tr>
<td>Chemicals</td>
<td>16,028</td>
<td>1.2%</td>
<td>110</td>
<td>0.4%</td>
</tr>
<tr>
<td>ICT &amp; Precision Instruments</td>
<td>8,393</td>
<td>0.6%</td>
<td>138</td>
<td>0.4%</td>
</tr>
<tr>
<td>Automotive</td>
<td>6,655</td>
<td>0.5%</td>
<td>133</td>
<td>0.4%</td>
</tr>
<tr>
<td>Aerospace</td>
<td>5,610</td>
<td>0.4%</td>
<td>112</td>
<td>0.4%</td>
</tr>
<tr>
<td>Machinery, Electrical &amp; Transport Equipment</td>
<td>22,748</td>
<td>1.7%</td>
<td>412</td>
<td>1.3%</td>
</tr>
<tr>
<td>Pharmaceuticals</td>
<td>10,023</td>
<td>0.7%</td>
<td>38</td>
<td>0.1%</td>
</tr>
<tr>
<td>Agriculture, Forestry &amp; Fishing</td>
<td>9,122</td>
<td>0.7%</td>
<td>409</td>
<td>1.3%</td>
</tr>
<tr>
<td>Mining &amp; Quarrying</td>
<td>39,646</td>
<td>2.9%</td>
<td>61</td>
<td>0.2%</td>
</tr>
<tr>
<td>Utilities</td>
<td>37,762</td>
<td>2.7%</td>
<td>327</td>
<td>1.0%</td>
</tr>
<tr>
<td>Construction</td>
<td>91,681</td>
<td>6.7%</td>
<td>2,036</td>
<td>6.5%</td>
</tr>
<tr>
<td>Communications</td>
<td>23,028</td>
<td>1.7%</td>
<td>227</td>
<td>0.7%</td>
</tr>
<tr>
<td>Digital, Creative &amp; Information Services</td>
<td>61,821</td>
<td>4.5%</td>
<td>1,174</td>
<td>3.7%</td>
</tr>
<tr>
<td>Financial Services</td>
<td>128,830</td>
<td>9.4%</td>
<td>1,116</td>
<td>3.6%</td>
</tr>
<tr>
<td>Business Services</td>
<td>97,528</td>
<td>7.1%</td>
<td>2,235</td>
<td>7.1%</td>
</tr>
<tr>
<td>Research &amp; Development</td>
<td>4,280</td>
<td>0.3%</td>
<td>126</td>
<td>0.4%</td>
</tr>
<tr>
<td>Education</td>
<td>89,676</td>
<td>6.5%</td>
<td>2,722</td>
<td>8.7%</td>
</tr>
<tr>
<td>Knowledge Services</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hotels &amp; Restaurants</td>
<td>39,601</td>
<td>2.0%</td>
<td>1,990</td>
<td>6.3%</td>
</tr>
<tr>
<td>Retail</td>
<td>71,016</td>
<td>5.2%</td>
<td>3,070</td>
<td>9.0%</td>
</tr>
<tr>
<td>Transport, Storage &amp; Distribution</td>
<td>149,580</td>
<td>10.9%</td>
<td>3,183</td>
<td>10.1%</td>
</tr>
<tr>
<td>Real Estate</td>
<td>98,091</td>
<td>7.1%</td>
<td>417</td>
<td>1.3%</td>
</tr>
<tr>
<td>Administrative &amp; Support Services</td>
<td>65,509</td>
<td>4.8%</td>
<td>2,432</td>
<td>7.8%</td>
</tr>
<tr>
<td>Public Admin &amp; Defence</td>
<td>67,915</td>
<td>4.9%</td>
<td>1,654</td>
<td>5.3%</td>
</tr>
<tr>
<td>Health &amp; Social Care</td>
<td>106,766</td>
<td>7.8%</td>
<td>4,079</td>
<td>13.0%</td>
</tr>
<tr>
<td>Other Services</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole Economy</td>
<td>1,373,399</td>
<td></td>
<td>31,378</td>
<td></td>
</tr>
</tbody>
</table>

Source: BIS analysis of ONS data.