Big data - Is it an opportunity or challenge?

Pooja Aggarwal¹, Nehal Mittal², Dr.Sanjiwani Kumar³

¹Pooja Aggarwal, B.tech Computer Science with 2 years work experience at hCentive Technologies Pvt Ltd, currently pursuing PGDM International Business at KJ SIMSR.
²Nehal Mittal, B.tech Electronics and Communication with 2 years work experience at Ericsson India Global Services Pvt Ltd, currently pursuing PGDM International Business at KJ SIMSR.
³Faculty, Quantitative Techniques and Research, KJ Somaiya Institute of Management Studies and Research, Mumbai

Abstract:
Objective: Big Data has gained huge popularity over the last few years among IT professionals and college professors etc. The term “big data” often refers simply to the use of predictive analytics, user behaviour analytics, or certain other advanced data analytics methods that extract value from data, and seldom to a particular size of data set. Analysis of data sets can find new correlations to “spot business trends, prevent diseases, combat crime and so on. The objective of this study is to explore the opportunities of using Big data and creating new opportunities for growing businesses in different sectors. Decisions that previously were based on Guess work, or on painstakingly constructed models of reality, can now be made based on the data itself. The finding of the Big Data helps in making decisions for the organization(IT, Healthcare, Banking, Retail, Pharmaceuticals sectors, etc.) for their betterment.

Methodology: The research was conducted using secondary data. Qualitative data was collected from various e-commerce websites like Amazon, Flipkart, Snapdeal where review of various customers about the Samsung smartphone are collected. As all customers were free to write any review, data was open ended and unstructured.

Limitations: The research is based on reviews of customers on e-commerce websites. It does not cater to customers who purchase Samsung smartphones from physical stores. As the data is huge, it becomes difficult to capture, store and visualise each data point.

Keywords: Big data, objective, secondary, Qualitative

I. Introduction

Nowadays, the world is moving towards the information society, where large amount of data are needed to extract better knowledge called Knowledge Discovery from Data, or KDD. It was in the early 21st century when the world first heard about the concept of big data and the attributes like too large, too unstructured and too fast moving were used for describing the nature of the data. The growing popularity of Big Data is mainly due to more number of devices and apps being used such as personal computers, mobile phones, government records, healthcare records, social media, climate sensors, airport terminals, etc. Business analytics is useful for a business to examine patterns and trends in large data sets. Examining the data helps a business generate models for future predictions of patterns and trends. Characteristics of big data have been defined by volume, velocity, variety and value. Volume implies the amount of data. While volume indicates more data, it is the granular nature of the data that is unique. For some organizations, this might be tens of terabytes, for others it may be hundreds of petabytes. Velocity is the fast rate at which data is received and perhaps acted upon. Variety indicated new unstructured data types. Unstructured and semi-structured data types, such as text, audio, and video require additional processing to both derive meaning and the supporting metadata. Along with these, data has intrinsic value, but it must be discovered. There are a range of quantitative and investigative techniques to derive value from data like discovering a consumer preference or sentiment. The importance of big data doesn’t revolve around how much data we have, but what we do with it. An organisation can take data from any source and analyze it to find answers that enable cost reduction, time reduction, smart decision making, new product development and optimised offerings. The purpose of this paper is to briefly describe the nature of big data, highlight its importance in the business world, and make the case for incorporating big data analytics as an essential tool in business. Through this research, we want to understand perception of customers towards Samsung smartphone through Content Analysis (Qualitative Research) of Customer Reviews (Secondary Data) posted on various websites. This research can be helpful is recognising shortcoming in smartphone and thus help in smartphone enhancement or new product development.
II. Literature Review

Sahil R. Kalra, Aarati Mahajan (2015), Big Challenges, Big Data and Big Data Opportunities revealed that the era of Big Data has just begun. Big data is used to describe exponential growth and availability of data, both structured and unstructured. Three V’s of Big data are Volume, Velocity and Variety. Big data helps to do accurate analysis. Some of the other goals behind using Big data are Cost reductions, time reductions, New Product development and optimized offerings and smarter business decision making. Some techniques associated with Big data analysis include Hadoop, Stratosphere, Calpont etc. Also, Big data is being called “The next frontier for innovation, competition and productivity.” This will give potential for making faster advances in many scientific disciplines and improving the profitability and success of many enterprises.

Laura B. Stokes, Joseph W. Rogers Et al (2016), Implications of Big data for Health System Pharmacy provided a perspective on how Big data can be applied to health system pharmacy. It also describes an approach for pharmacy leaders to effectively use Big data. Big data can also enhance the development of patient-centred pharmacy services as health care information expands and become more integrated. Patients may utilize these data to empower their health care decisions, whereas providers may use Big data to obtain real time information about their patients or for decision support. The impact of Big data on the Pharmaceutical industry is basically to ensure that the right patient has the right medication at the right time.

Hasan Al-Sakran (2014), B2C E-Commerce Fact-Based Negotiation Using Big Data Analytics and Agent-Based Technologies focused on application of intelligent agent in negotiation between buyer and seller in B2C Commerce using big data analytics. Big data analytics also creates new opportunities for Bidding. Using big data analytics, sellers may learn to predict the buyers’ negotiation strategy and therefore adopt optimal tactics to pursue results that are to their best interests. The goal of negotiation is to seek a solution that optimizes utility value for both of them. Experimental design is used to collect intelligent data. Such approach will improve quality of negotiating decisions for both parties. The model proposed in this paper allows negotiators to engage in multi-parties negotiations. From that analytics knowledge, they may get better with selecting and achieving goals and taking correct actions.

Tanvir Ahammad, Md. Sajib Al Mamun Et al (2016), B2C Towards the Application of Big Data: A New Way to make Data Driven Healthcare Decision emphasized on big data application and demonstrated that the analysis of healthcare data paves the way to make a better healthcare decision, reduce cost, and raise healthcare consciousness. The experimental analysis has been accomplished with a proposed methodology based on existing Text Mining and Natural Language Processing (NLP) techniques. Here, healthcare information is first collected from different text data sources, e.g., healthcare news, blogs, social media and patient reports. This information is then stored in DB in a structured form for Data mining. This data is open source, where one patient can view others discussions who have the same disease or condition, track and share their own experiences and so on.

Regina Henry, Santosh Venkatraman (2015), Big Data Analytics the next big learning opportunity examined the rapidly growing field of Big Data Analytics and studied why and how big data analytics needs to be integrated into business skill sets and curriculum designs. The research will provide a practical framework to design and teach the skills sets needed to solve organizational problems by analysing the vast amounts of data that are being generated and stored. This research focuses on how undergraduate business schools may help students in higher education gain the big data and data analytics skills and experience necessary to fill the current employment gap of trained professionals in the field. In this, the whole idea of Big data is to sift through the mountains of data, and get actionable insights.

Tim J. Gabel, Cathy Tokarski (2014), Big data and Organization design: Key challenges await the Survey Research Firm aimed to share a practitioner’s perspective on the challenges of restructuring a knowledge-worker company in the midst of the big data revolution. These challenges include retooling fundamental human resource processes such as recruiting and hiring, performance management, and talent development. The case of RTI International, a non-profit U.S Organization, is taken and discussed. For RTI to become a stronger organization in the future, it must successfully address three large challenges posed by the big data phenomenon – Seize Opportunities, Survive Threats and Define what client’s need – and don’t need. After meeting these challenges, RTI will usher in significant changes both to their organization and the clients they serve. They recognize the urgency of capturing and formulating insights from big data at a time when they can enhance optimal decision making, both for their organization and for their clients.

Sergei V. Kalinin, Bobby G. Sumpter Et al. (2015), Big – Deep – Smart data in imaging for guiding materials design discussed the new opportunities in materials design enabled by the availability of big data in imaging and data analytics approaches, including their limitations, in material systems of practical interest. Such methodologies are particularly appropriate to explore in light of continued improvements in atomistic imaging, modelling and data analytics methods. The proper use of big data, or the vast amount of data that can be measured and simulated, can act as a bridge between theory and functional imaging. This vision could be
facilitated by broadening the scope of academic education, developing new university curricula and degrees in undergraduate and graduate studies, which combine statistical learning, programming, physical chemistry, and materials science, and highlighting the intrinsic links between disciplines (for example, statistical physics and information theory).

Poonam Sawant, B.L. Desai (2015), Big data analytics to predict Customer Behavior for Personalized Banking Services found that as customer interactions in banking move from in-person to digital channels, you not only have to react faster; you must also be able to predict future behavior which is very useful to achieve better operational efficiency and create a new product market. It is possible by using high end predictive analytics and Big Data analytics is a better solution. This paper represents overview, current scenario, challenges and importance of big data and big data analytics in banking sector to forecast its future scenario. Big data and Hadoop can enable bankers to connect with customers through multiple channels by harnessing the massive volumes of new data available today. Big data provides organizations with greater opportunities by exposing customer’s hidden behavioural patterns and helps to bridge gap between what customers want to do and what they actually do. This information is useful to make business decisions and improve services to increase operational efficiency and create new product or markets.

Hwan – Seok Yang (2016), The study on the Business Development Prospects Analysis using Big Data analyzed applied technology for development model of Big data business through it. In particular, they divided the development model for Big data business step-by-step and analyzed consideration about this systematically. The companies and public institutions based on the analyzed result in this paper will provide advantages to production of big data systems. In this paper, they proposed development model for big data business in order to lead to creation of new service in data, service, and infra and pull activation of big data market.

Lidong Wang, Cheryl Ann Alexander (2015), Big Data In Design And Manufacturing Engineering introduced Big Data, its characteristics and a number of issues of Big Data in design and manufacturing engineering. These issues include design and manufacturing data, Big Data benefits and impacts and its applications and opportunities in Automotive Industry, Semiconductor Manufacturing and Integrated Circuits, Missile Plant, Cloud-based design and manufacturing, Medical device design, Quality Assurance and Logistics, CAD/CAE/CAM and CAD. Methods, technologies and some technology progress around Big Data are presented in this study. Big Data offering benefits like- Defect tracking and product quality, Improvements in supply planning, Improved product manufacturing processes, Driven efficiency across the extended enterprise, improved service. Charles G. Jobs, Steven M. Aukers Et al. (2014), The Impact Of Big Data On Your Firms Marketing Communications: A Framework For Understanding The Emerging Marketing Analytics Industry introduced how Big data is driving the adoption of broader and increasingly sophisticated quantitative analysis technique across media channels by large, medium and even smaller sized firms. It also provide insights into qualitative opportunities and challenges that marketing organizations face operating in a big data world. In here the Big Data is defined as a cultural, technological, and scholarly phenomenon that rests on the interplay of 3 factors (Technology, Analysis and Mythology). This paper unveils specifics about the ecosystem of firms in the big data arena as it relates to marketing communications and advertising.

Michele De Gennaro, Elena Paffumi Et al (2016), Big Data for Supporting Low-Carbon Road Transport Policies in Europe: Applications, Challenges and Opportunities provided the scientific community with a comprehensive overview of the applications of a data processing platform designed to harness the potential of big data in the field of road transport policies in Europe. This platform relies on datasets of driving and mobility patterns collected by means of navigation systems. Luis M. A. Bettencourt (2013), The Uses of Big Data in Cities explored how big data can be useful in urban planning by formalizing the planning process as a general computational problem. As such the primary role of big data in cities is to facilitate information flows and mechanisms of learning and coordination by heterogeneous individuals. However, processes of self-organization in cities, as well as of service improvement and expansion, must rely on general principles that enforce necessary conditions for cities to operate and evolve. Aimia Coalition (2014), How loyalty programs and Big Data Analytics are facilitating CRM in the retail sector found that Big data analytics is used in a way with retailers to make decisions on how they stock their store. For example, through analysis of customers’ Nectar card data, in addition to the purchasing habits of non-Nectar card users, UK supermarket chain Sainsbury’s discovered Grape-Nuts cereal was worth stocking, despite poor sales. This was due to the fact that those that did purchase it were found to be extremely loyal and were often big spenders. Sacrificing shelf space in order to retain and attract big spending customers is a small price to pay for a retailer, and is a connection that they would be unlikely to discover without big data analytics. With the advent of big data analytics, customer loyalty programs are now facilitating the implementation of CRM in ways only restricted by an analyst's imagination. Miklos A. Vasarhelyi, Alexander Kogan Et al. (2012) Big Data in Accounting: An Overview focused on the sources and uses of Big Data in accounting (measurement) and auditing (assurance). The paper also discusses the interaction of Big Data and traditional sources of data, as well as Big Data’s impact on audit
judgment and behavioural research. Both accounting academics and accounting practitioners will benefit from learning about the significant potential benefits of Big Data and the inevitable challenges and obstacles in the way of its utilization.

Effect of Big data on various sectors

<table>
<thead>
<tr>
<th>BUSINESS SECTOR</th>
<th>EFFECTS OF BIG DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>BANKING</td>
<td>With large amounts of information streaming in from countless sources, banks are faced with finding new and innovative ways to manage big data. While it’s important to understand customers and boost their satisfaction, it’s equally important to minimize risk and fraud while maintaining regulatory compliance. Big data brings big insights, but it also requires financial institutions to stay one step ahead of the game with advanced analytics.</td>
</tr>
<tr>
<td>EDUCATION</td>
<td>Educators armed with data-driven insight can make a significant impact on school systems, students and curriculums. By analyzing big data, they can identify at-risk students, make sure students are making adequate progress, and can implement a better system for evaluation and support of teachers and principals.</td>
</tr>
<tr>
<td>GOVERNMENT</td>
<td>When government agencies are able to harness and apply analytics to their big data, they gain significant ground when it comes to managing utilities, running agencies, dealing with traffic congestion or preventing crime. But while there are many advantages to big data, governments must also address issues of transparency and privacy.</td>
</tr>
<tr>
<td>HEATHCARE</td>
<td>Patient records. Treatment plans. Prescription information. When it comes to health care, everything needs to be done quickly, accurately – and, in some cases, with enough transparency to satisfy stringent industry regulations. When big data is managed effectively, health care providers can uncover hidden insights that improve patient care.</td>
</tr>
<tr>
<td>MANUFACTURING</td>
<td>Armed with insight that big data can provide, manufacturers can boost quality and output while minimizing waste – processes that are key in today’s highly competitive market. More and more manufacturers are working in an analytics-based culture, which means they can solve problems faster and make more agile business decisions.</td>
</tr>
<tr>
<td>RETAIL</td>
<td>Customer relationship building is critical to the retail industry – and the best way to manage that is to manage big data. Retailers need to know the best way to market to customers, the most effective way to handle transactions, and the most strategic way to bring back lapsed business. Big data remains at the heart of all those things.</td>
</tr>
</tbody>
</table>

III. Research Methodology

The research design used here is content analysis (Qualitative study) of secondary data. The secondary data here refers to the reviews posted by users of Samsung smart phone users on various e-commerce websites like Amazon, Flipkart, Snapdeal etc. Perception of customers towards the Smartphone was examined by performing content analysis of about 2000 recommendations posted on Amazon, Flipkart etc by customers themselves. Some strategic keywords were identified and the frequencies of their occurrence were analysed to study the qualitative aspect of the Samsung smart phone. The reviews were unstructured in nature, they had to be converted into some meaningful structured form of information, so that some proper analysis could be carried out on the study. The unstructured data was converted to structured data by grouping it into two categories. Since, the reviews that were posted on the websites were either on the positive or negative, they were grouped into these two basic categories to make them more structured.
IV. Qualitative Data Analysis

People judge smartphone on various parameters with camera, RAM, looks, Battery, Price being most important. There are certain keywords which are used extensively by reviewers in order to define various aspects of a Smartphone. Few most commonly words along with their counts are depicted in the chart below. MS Excel was used to perform count method over a set of 2000 data points to find the occurrence of parameters used by customers to give their reviews.

![Frequency of various parameters](chart.png)

We see dual sim for defining the type of a phone and mAh which is a parameter of the battery of the phone are being used to a greater extent. Pixel and Android are also among common words being used. The reviews were evaluated and then some specific keywords were chosen for evaluating these reviews. There are certain parameters which are important to people while buying the Smartphone. Few of them are camera, battery, looks and price. Most of words which describe these factors along with the frequency of their occurrence are analysed. Frequency of occurrence of these keywords was calculated with the help of Microsoft Excel.

i) **Camera**: It is one of the most important judgement parameter for a smartphone. Depending upon their perspective people use various words for defining the camera of phone. They use both positive and negative words for describing Camera of smart phone. We found the count of these positive and negative words and thus prepared frequency table as below.

<table>
<thead>
<tr>
<th>Positive Keyword</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Great</td>
<td>45</td>
</tr>
<tr>
<td>Good</td>
<td>92</td>
</tr>
<tr>
<td>Amazing</td>
<td>14</td>
</tr>
<tr>
<td>clear</td>
<td>19</td>
</tr>
<tr>
<td>satisfied</td>
<td>5</td>
</tr>
<tr>
<td>Terrific</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Negative Keyword</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bad</td>
<td>23</td>
</tr>
<tr>
<td>Unclear</td>
<td>6</td>
</tr>
<tr>
<td>Dissapointed</td>
<td>15</td>
</tr>
<tr>
<td>Pathetic</td>
<td>3</td>
</tr>
<tr>
<td>Poor</td>
<td>34</td>
</tr>
<tr>
<td>Terrible</td>
<td>1</td>
</tr>
<tr>
<td>Worst</td>
<td>9</td>
</tr>
</tbody>
</table>
Among various words, “Good” is being used by people to the maximum extent which indicates that people are quite satisfied with the performance of camera of the phone. Term “Great” is also used to a large extent. Among the negative words Poor and Bad are used for most of the times.

ii) **Battery**: It is also an important judgement parameter while selecting a smart phone. There are people who need phone of greater battery life as compared to any other factors. People use both positive and negative words for describing battery life and performance of Samsung smart phone. We found the count of these positive and negative words via Excel and thus prepared frequency table as below, thus prepared word map to depict the same.

![Word Map for Camera Reviews](image1)

![Word Map for Battery Reviews](image2)
We see “Discharge” is the word which is being used most extensively by reviewers. This is sort of negative word. People don’t seem much impressed by the battery life. Though few people have used words like “Good” in order to describe the battery.

\[\text{iii) Looks of the smart phone:}\text{ It is also an important judgement parameter while selecting a smart phone. This is one of the first things which determine perception of people. Looks include everything from phone color, weight, width and appearance of smartphone. Many positive and negative words are used for describing the Smartphone. We found the count of these positive and negative words via Excel, prepared frequency table as below, thus prepared word map to depict the same.}\]

![Word Map](image)

\[(\text{Fig 5: Reviews about Looks – Greater font size indicates greater frequency})\]

Among various words “Great” is being used by people to the maximum extent which indicates that people are quite satisfied with the looks of the phone. Good is also used to a large extent. A prominent positive word “Sleek” is also being used. Among the negative words, “Heavy” and “Poor” are used for most of the times. This indicates that more or less, customers who purchased Samsung smartphone are satisfied by its appearance, its look and feel.

\[\text{iv) Price: It is the filter parameter of people, everyone buy phone according to the budget they have. Mostly people think product as if it has value for money or not which makes it worth for the price they are paying or not. There are very less words people have used to describe it.}\]

![Price Map](image)
Big data - Is it an opportunity or challenge?

We see “Low” and “Costly” are the words which are being used most extensively by reviewers. This is sort mixed opinion. Few people think the phone is low priced as to the functionality and few think it is priced at higher price comparatively.

**Further Analysis**

After analysing the frequency of the keywords, we went about mapping these keywords to the context in which they were used. For example, we can see that the word “Good” has appeared in the analysis 40 times. We then analysed the extent to which user was compelled to use this word in the review. Following are the results that we got:

<table>
<thead>
<tr>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td>Looks are good, Battery is good</td>
</tr>
<tr>
<td>Great</td>
<td>Great looks, Great camera</td>
</tr>
<tr>
<td>Average</td>
<td>Battery life is average, average looks</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>V. Conclusion and Recommendations</td>
<td></td>
</tr>
</tbody>
</table>

**5.1 Conclusion**

The era of Big Data has started, in which Data deluge will keep on increasing throughout the years, and each data scientist will have to handle much more quantity of data every year. This data is becoming more diverse, larger, and faster. Though better analysis of the large volumes of data that are becoming available, there is the potential for making faster advances in many scientific disciplines and improving the profitability and success of many enterprises. Those who will benefit from their ‘Data Clients’ capital, will open new perspectives towards greater competitiveness and innovation. In this document, we discussed a brief overview on the Big data topic, including the main concerns for the future, as well as costumer management as an application of big data. The more data will increase in Big Data universe, the more potentialities will be created; hence requires more analysis. Big Data will allow us to extract insights that no one has extracted before. However, it is still under development and current approaches and tools are very limited to deal with the new real Big Data requirements.

From the research, we found that the count of positive reviews is higher than that of negative reviews.

So it can be concluded that Samsung smartphone has been able to impact on a highly positive note in the minds of the consumer and end users. Mapping of the keywords led to the conclusion that the customers who purchased Samsung smartphone are satisfied and happy with the look and feel of the smartphone and also like camera quality of the smartphone. However, customers are not satisfied with the battery life as they complain it discharges quickly and the price of the smartphone has more or less equal positive and negative reviews.

**5.2 Recommendations**

As it can be seen from the reviews against battery life, customers are not completely satisfied with the Samsung Smartphone. So in this volatile, uncertain, complex and ambiguous world, the brand image has to be maintained. Samsung has to take care of the fact that if no. of negative reviews increases, it may hamper its brand image and thus lose its market share. Samsung has to be careful about the count of positive and negative reviews, it should have at least 1 positive review among every 2-3 negative reviews as reviews play very important role in building customer perception.

This is biggest challenge in big data as open ended, unstructured, qualitative data analysis is very tedious task. Text mining model is effective in such analysis and is recommended to sort out new consumer
problems. But as it is based on assumption that consumer is clear about the product he wants to buy which is again another challenge. Thus we would recommend Samsung to take immediate action for any negative reviews about its Smartphone on e-commerce websites in order to achieve full customer satisfaction.

Limitations
1. The sample is open ended and unstructured, thus does not have any demographic categorisation.
2. Visualisation of data is tedious and not completely accurate.

References
[1]. Aimia Coalition (2014), How loyalty programs and Big Data Analytics are facilitating CRM in the retail sector
[2]. Charles G. Jobs , Steven M. Aukers Et al. (2014), The Impact Of Big Data On Your Firms Marketing Communications: A Framework For Understanding The Emerging Marketing Analytics Industry
[3]. Hasan Al-Sakran (2014), B2C E-Commerce Fact-Based Negotiation Using Big Data Analytics and Agent-Based Technologies
[5]. Laura B. Stokes, Joseph W. Rogers Et al (2016), Implications of Big data for Health System Pharmacy
[6]. Lidong Wang, Cheryl Ann Alexander (2015), Big Data In Design And Manufacturing Engineering
[7]. Luis M. A. Bettencourt (2013), The Uses of Big Data in Cities
[8]. Michele De Gennaro, Elena Paffumi Et al (2016), Big Data for Supporting Low-Carbon Road Transport Policies in Europe: Applications, Challenges and Opportunities
[12]. Poonam Sawant, B.L. Desai (2015), Big data analytics to predict Customer Behavior for Personalized Banking Services
[13]. Regina Henry, Santosh Venkatraman (2015), Big Data Analytics the next big learning opportunity
[14]. Sahil R. Kalra, Aarati Mahajan(2015), Big Challenges, Big Data and Big Data Opportunities
[15]. Sergei V. Kalinin, Bobby G. Sumpter Et al. (2015), Big – Deep – Smart data in imaging for guiding materials design
[17]. Tim J. Gabel, Cathy Tokarski (2014), Big data and Organization design: Key challenges await the Survey Research Firm
[18]. Text books referred
[19]. Kevin Lane Keller (2011), Marketing Research
[20]. Zikmund(2013), Marketing Research
[21]. Naresh K. Malhotra (2009), Marketing Research, An Applied Orientation
[22]. Anderson(2007), Marketing Research