CONSUMER INSIGHT AS COMPETITIVE ADVANTAGE
USING BIG DATA AND ANALYTICS

Adnan Veysel Ertemel, (PhD)

Istanbul Commerce University, Turkey

Abstract:

Digital revolution serves as a competitive advantage to businesses that are able to analyze consumer behavior in order to gain insights for their strategic advantage. After the advent of Internet, the past two decades witnessed generation of vast amount of business data. The amount of data is so huge that traditional database management system approaches falls short of managing and analyzing this data. This paper explores the characteristics of this phenomenon called Big Data together with Analytics as a tool for marketers to gain insights about consumer behavior and hence provide competitive advantage to the businesses. It also discusses some best practices as case studies.

Keywords: *Big Data, Big Data Analytics, Consumer Insight, Social Media, Sentiment Analysis*

1. Introduction:

Data provide behavioral insights about consumers. In today's highly competitive environment, translating those insights into market advantage is a key differentiator as described by Fayyad et al. (1996). In recent years, businesses generate more data than they are able to use. Now, the marketers are not only capable of getting rich data on consumer behavior, they can also do this in real time. Formerly possible but not feasible idea of capturing and analyzing in depth data about individual customers now also became feasible. This is because the rapid developments in technology enabled widely available data, which is much cheaper to access and store as stated by McAfee et al (2012)

Analytics generally refers to the tools that help find hidden patterns in data. What is different today is the enormous volume, velocity, and variety of data available about individual consumers. This results in a Big Data revolution that has the potential to lead to entirely new ways of understanding consumer behavior and hence help formulate new marketing strategies. In this paper, Big Data analytics is defined as the extraction of hidden insight about consumer behavior from Big Data applying that insight in business decisions. Big Data is considered a new form of capital in today's marketplace as described by Mayer-Schönberger and Cukier (2013).

This paper first examines the Big Data and Big Data Analytics phenomenon as a strategic weapon for the marketer to gain insights on consumer behavior. Data, Big Data, and Big Data Analytics concepts are investigated first. Afterwards, some of the best practices of the field are discussed.

2. Literature Review

2.1. From Data to Insights

Data is fundamental to information and information is fundamental to insights. These concepts are sometimes mistaken with each other. The concepts are explained below.

- Data: Data can be defined as the raw numbers captured to measure something according to some agreed to standards. Having consistent standards is critical in this definition. It's also important to note that data is raw. It has no meaning on its own. It could be the IP address of a visitor, or for weather reports, temperature, air pressure, humidity etc.

Figure 1: From Data to Insights

Source: http://sageassessments.blogspot.com.tr/
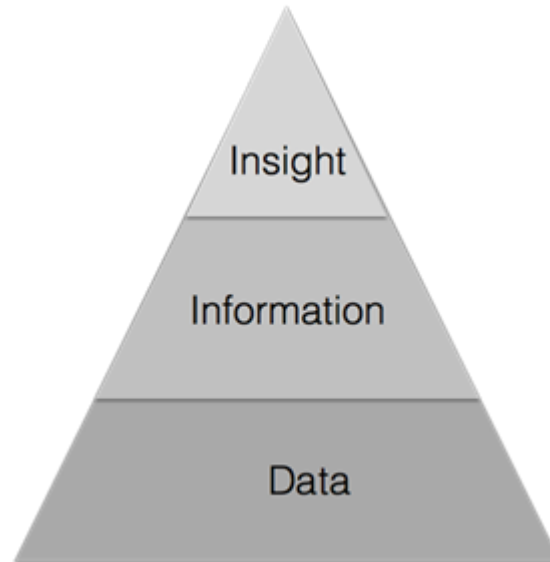
- Information: Information refers to the sum of data points used to understand something about what's being measured. Collectively all data points give some very useful information. In the end, information is there to interpret the data and give it a meaning.

- Insight: Insight is gained by analyzing data and information to understand what is going on with the particular situation or phenomena. . Insight is acquired when conclusions are drawn from data and information. Insight building is the ultimate goal for data analysis and the world of big data. Insight serves as a means to make more informed and better business decisions.

As an illustration of the concept in weather situation, knowing the average temperature for your location might be helpful, but knowing that the average is steadily increasing over five years gives a better understanding and puts the information into context. And the context is relevant to the business problem at hand.

2.2 . Big Data

The estimate size of the digital universe in 2013 was 4.4 zettabytes according to Cisco (2014) (1 zettabyte is equivalent to 250 billion DVDs). By 2020, IDC group (2014) predicts that the digital universe is expected to reach 44 zettabytes. Big Data volume, in other words, the market for storing and analyzing Big Data is expected to double in size every two years. This data includes but is not limited to emails, audios, videos, images, logs, posts, search queries, health records, online transactions, social networking interactions, scientific data, sensor obtained data in smart devices and smartphone generated data, Eaton et all (2012). The massive growth of data made it impossible to manage it with conventional Database Management Systems (DBMSs). This leads to the definition Big Data. Gartner (2013) defines big data as high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision-making, and process automation. Big data can be characterized with three distinct dimensions; high volume, high velocity and high variety.

Figure 2: Big Data Characteristics

Source: http://hadooptrainingsinhyderabad.blogspot.com.tr
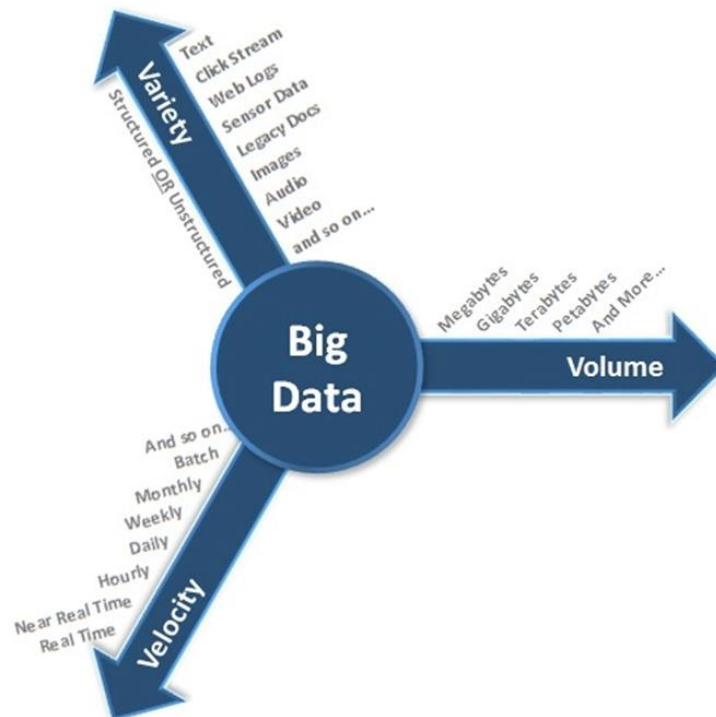
- Volume: More than ninety percent of data created in history was created in the past two years. By 2020, the data generated is going to be 50 times more than the amount of data in 2011. With conventional methods, the creation of so much data used to cause serious problems. But now, with decreasing storage costs such big data this is no longer a problem to the marketer.
- Velocity: Data Velocity refers to the speed at which the data is created, stored and analyzed. In the past, big servers required substantial time to process data. In the big data concept, data is created almost in real-time. The new phenomenon, Internet of Things takes this one step further with machines sending their data at the moment of creation. Data is created at an enormous speed. As an example, as of year 2015, every minute, 100 hours of video is uploaded to YouTube, and 20 million photos are viewed.
- Variety: Many sources of Big Data provide a diverse richness that far surpasses traditional data from the past. A major difference between contemporary Big Data and traditional data, according to Integreon Insight (2012) is the shift from structured transactional data to unstructured behavioral data. Structured data like scanner or sensor data, records, files and databases have been collected by marketers for some time. Unstructured data include textual data like from blogs and text messages and non-textual data like from videos, images, and audio recordings. Much unstructured data are captured through social media, where individuals share personal and behavioral information with friends and family. Semi-structured data incorporate various types of software that can bring order to the unstructured data.

The ever-increasing amounts of Big Data lead to the question of value. The task is to eliminate unimportant and irrelevant data, so that the remaining data are useful. According to Lycett (2013), the remaining pertinent data also needs to be valuable for obtaining insight and domain-specific interpretation. The challenge, according to Oracle (2012) is to identify what is relevant and then rapidly extracts that data for timely analysis.

2.3. Big Data Analytics

Big data analytics is classified as descriptive, predictive and prescriptive analytics.

- Descriptive Analytics: Descriptive analytics aims at counting data and summarizing it with understandable reports. For example, with descriptive analytics, the impact of an advertising campaign can be analyzed. Additionally, the hidden patterns, trends in consumer behavior etc. can be spotted with descriptive analytics tools.

- Predictive Analytics: Predictive analytics uses historical data to detect recurring patterns and make trajectory about the future by making use of this historical data. Weather forecasts are a typical example to predictive analytics phenomenon. to infer data that we either can't collect, didn't collect or haven't collect yet.

- Prescriptive Analytics: Prescriptive analytics uses simulation modeling, optimization, what-if analysis, times series models and the like techniques to identify the best alternatives to minimize or maximize some objective. It's being used in many areas. For instance, it can be used by to determine the best pricing and advertising strategy for revenue maximization. The statistical methods for predictive analytics can be combined with optimization to make decisions assessing the uncertainty in the data.

3. Using Big Data and Analytics as a Competitive Advantage:

For centuries, the most common practice in business management used to be making strategic business decisions based on Highest Paid Person's Opinion (HiPPo) in the company as stated by McAfee and Brynjolfsson (2012). This intuition-based approach was not an option but a necessity due to the lack of an alternative.

In the last decade, thanks to the technical advances in consumer analytics and big data phenomenon, marketing has become much more measurable and technical than ever before. Marketers are now able to afford storing and analyzing vast amount of big data about their customers coming from a variety of sources to get 360-degree view of customer behavior. A lot of insight could now be obtained out of analyzing such big data. According to Athey (2013), there are now companies, which throw away the opinions of their senior executives and started to build genius teams of data scientists making sense of customer data. The valuable insights and conclusions these teams make give the direction as to where should the organization should be heading to.

Past research undertaken by Lavalle et al (2013) shows that many top performing organizations use analytics as a key differentiator The past research has also found that top-performing organizations made decisions based on this approach double the rate of lower-performing organizations. In another study by, McAfee and Brynjolfsson (2012), it's concluded that the more companies characterized themselves as data-driven, the better they performed on objective measures of financial and operational results top third of their industry in the use of data-driven decision making were on average, 5% more productive and 6% more profitable than their competitors.

3.1 New Channels, New Data, New Insights:

The social media revolution opens up brand new channels that enable more thorough view of consumer behavior analysis and insights. With foursquare-like services, marketers gain more understanding on consumer geospatial data whereas mere text messages, when aggregated, can give a tone of emotional reactions of consumers to a specific phenomenon.

A Hypothetical Scenario
The following scenario can explain, in a broader sense, the potential and capabilities of big data concept:
Imagine you are a blogger with thousands of followers. You decided to buy a new tablet from the online retailer, Amazon.com. When you unbundle the box and try to open your brand new tablet, you discover that it came broken.

As a result of this bad experience, you sat on your table and write a very negative blog revealing your experience. Imagine a call agent from amazon calling you to apologize and correct their mistakes by sending a brand new tablet to your address and saying: 'When we looked at your LinkedIn profile we saw that you changed your job recently, when we looked at your foursquare check-ins, it seems that you're on holiday. So, where do you want us to deliver your new tablet?'

This interesting scenario illustrates the promise and potential of big data: it doesn't only imply the data the businesses own about their customers, but it also implies the data that's out on the internet about those customers. As data is available in more varieties, more valuable insights could be found.

4. Big Data Case Studies

4.1. Sentiment Analysis and Brand Associations – Nike Example

Sentiment analysis is an emerging domain regarding social media analytics according to Rajani (2009). It analyzes raw text data within user generated content to look at jargon, sentiment, orientation, polarity, label and the like and collapse the tone of the individual message down to 1 dimension, whether the message is positive or negative. By analyzing the degree of satisfaction of individual stakeholder, corrective action can be taken based on priorities, if necessary. A combination of keyword specific classification approaches as well as semantic classification approaches such as linguistic meaning may be used to achieve this goal. Such methods open up the possibility to address customer concerns in general. Taking this one step further, some advanced software packages even produces fairly rich brand association maps, as a result of analysis of all publicly available data accumulated since year 2007.

As an illustration, below is a Brand Associations map showing associations the public society attribute to Nike brand based on tens of thousands of tweets, public posts etc.

Figure 3: Nike Brand Associations Map
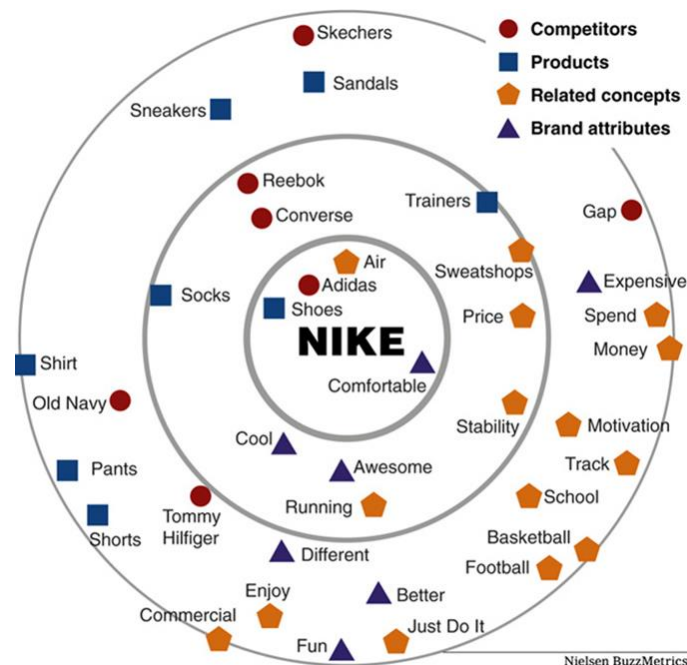
Source Nielsen, 2009

With the valuable insights brand associations map provide, the brand can;

- Understand, in order, how a category relates to a brand (i.e. for Nike, Shoes, Trainers, Socks, Shirts etc.)
- See how consumers see as the brands' closest competitor and other competitors with relative importance
- See related concepts and illuminate popular markets with great potential
- Identify early warning signs to potential threats and
- Recognize messaging attributes that can resonate to consumers in advertising efforts

4.2 Netflix

As an online streaming movie provider platform, with more than 60 million worldwide subscribers and more than 30 million daily movie plays, Netflix has got a very large data on consumer viewership habits. It doesn't only look at apparent rating data, but it looks more deeply on customer behavior data including; when consumers pause, rewind, or fast-forward, what day they watch content. The date they watch. What time they watch content. Where they watch (zip code). What device they use? When pause is pressed and content left and if they ever come back again., the ratings given, searches browsing behavior, and even data within movies by taking various screenshots to look at "in the moment" characteristics (Bulygo, 2013).

After analyzing this wealthy data, Netflix was pretty confident to say what content is ideal to hit the mainstream market and they went on to make a TV-series production from scratch. This production named House of Cards became a real success, bringing an estimate of 17 million new subscribers. Even in promoting the content to different audiences, Netflix prepared ten different versions of House of Cards trailer. For instance, the starring actor Kevin Spacey fans watched the trailer focusing on him whereas people who liked female oriented movies saw a different variation of the content. Netflix is also very successful with its recommendation engine. Based on its deep understanding of each and every customer, it's able to suggest the content that's most likely to interest the individual audience. Therefore, Netflix earns two thirds of its revenue attributing to recommendation engine Carr (2013)

4.3 Zest Finance

As a payday lender, Zest Finance is an Internet based financial company. Since the foundation in 2009, it received more than 100.000 loan applications. It makes heavy use of data to augment traditional underwriting. Using a regression model with thousands of variables and more than 10 models it tries to come up with insights on whom to depend more and whom to avoid as borrowers. By analyzing data, it found out that contrary to the normal way of filling the form where initials are in uppercase and other letters are in lowercase, those who do not plan to give back money fill in the initial application form all in capitals or all in lower case. After this finding, it began to discriminate and tend to reject those applications flagged as riskier. As a result of its data-driven innovative efforts, Zest Finance has achieved a %40 reduction in loan default rates and enjoyed a %25 increase in market share. Zest Finance also received interest of venture capitalists since 2011; it has attracted $62 million in venture financing, plus $50 million in debt financing as stated in BloombergBusiness (2014).

5. Conclusion

Marketers are increasingly using big data to gain invaluable insights about their customers. Big data has far reaching effects even for business management itself – businesses resort to big data to guide them as the wheat to do next in their strategic decisions, as opposed to relying on their CEO's opinion. This paper discussed big data and big data analytics concepts from the marketing and business management standpoint. Having examined the corresponding phenomenon, this paper examines the implications and potential of big data for the marketer. Finally, some of the industry best practices are discussed.

References

Ather, S., "How Big Data Changes Business Management", https://www.gsb.stanford.edu/insights/susan-athey-how-big-data-changes-business-management (accessed 21 November 2015)

BloombergBusiness(2014), "Lender Charging 390% Uses Data to Screen Out Deadbeats", http://www.bloomberg.com/news/articles/2014-10-01/lender-charging-390-uses-data-to-screen-out-deadbeats (accessed 20 November 2015)

Bulygo, Z.(2014), "How Netflix Uses Analytics To Select Movies, Create Content, and Make Multimillion Dollar Decisions", https://blog.kissmetrics.com/how-netflix-uses-analytics/ (accessed 20 November 2015)

Carr, D. (2013), "Giving Viewers What They Want", http://www.nytimes.com/2013/02/25/business/media/for-house-of-cards-using-big-data-to-guarantee-its-popularity.html (accessed 20 November 2015)

Cisco (2014). "Visual networking index IP traffic chart", http://www.cisco.com/cdc_content_elements/networking_solutions/service_provider/visual_networking_ip_traffic_chart.html (accessed November 18, 2015)

Eaton C., Deroos D., Deutsch T., Lapis G and P.C. Zikopoulos, (2012), "Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data", Mc Graw-Hill

Evans, James R., and Carl H. Lindner (2012), "Business analytics: the next frontier for decision sciences." Decision Line 43.2 4-6.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996), "From data mining to knowledge discovery in databases", AI Magazine, 17(3), p.37–53.

Gartner (2013), "Gartner IT Glossary - Big Data", http://www.gartner.com/it-glossary/big-data (accessed 15 November 2015)

IDC (2014). "The digital universe of opportunities: Rich data & the increasing value of the internet of things" http://www.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf   (accessed 19 November 2015)

Integreon Insight (2012). "Big just got bigger. Grail Research", 1–17 http://www.integreon.com/pdf/Blog/Grail-Research-Big-Data-Just-Got-Bigger_ 232.pdf (accessed 20 November 2015)

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2013). "Big data, analytics and the path from insights to value", MIT sloan management review, 21.

Lycett, M. (2013). "'Datafication': Making sense of (big) data in a complex world", European Journal of Information Systems, 22(4), 381–386.

Mayer-Schönberger, V., & Cukier, K. (2013). "Big data: A revolution that will transform how we live, work, and think", New York: Houghton Mifflin Harcourt.

Nilesen Brand Association Map (2009), http://www.nielsen.com/content/dam/nielsen/en_us/documents/pdf/Fact%20Sheets/Nielsen%20Brand%20Association%20Map%20-%20US.pdf (accessed 18 November 2015)

Oracle (2012). "Big data for the enterprise" Oracle White Paper, 1–14 http://www.oracle.com/us/products/database/big-data-for- enterprise-519135.pdf (accessed 20 November 2015)

Rajani, N. F. (2011). "Social media analytics". Published thesis, Birla Institute of Technology and Science, Pilani (Rajasthan).

Scheer, L. K., & Stern, L. W. (1992). "The effect of influence type and performance outcomes on attitude toward the influencer", Journal of Marketing Research, 29(1), 128-142.