**Big Data Technology Review**

**Big Data in Retail**

**Assignment 1**

**By:**

**Sanika Deshpande**

**Student ID : 805600277**

**CPSC 531-01 (13293)**

**Fall, 2016**

**Professor: Dr. Chun-l Phillip Chen**

**Department of Computer Science**

**California State University, Fullerton**

**Sept 27, 2016**

**Table of Contents**

Abstract .................................................................................................................................................. 4

1. Introduction.......................................................................................................................................5

1.1 What is Big Data?.................................................................................................................5

1.2 Why Big Data important?.....................................................................................................6

2.0 Literature Review ..............................................................................................................................7

2.1 Big data used in Retail .........................................................................................................7

2.2 Need for Retail Big Data Analytics.................................................................................... .8

2.3 Retail Application Process…………………………………………………………………8

2.4 Working of Big Data in Retail……………………………………………………………10

2.4.1 Data …………………………………………………………………………………10

2.4.2 Big data processing steps……………………………………………………………10

2.4.3 Working of Big data analytics………………………………………………………10

2.5 Five use cases in Retail…………………………………………………………………...10

2.6 Proposed Agent based big data analytics in retail ………………………………………..12

2.7 Retailer’s Biggest Obstacles to Success with Analytics………………………………….14

2.8 Big data’s Impact on Retail Process……………………………………………………....14

2.9 Retail Functions with gain benefit from Big data………………………………………...15

3.0 Discussion.......................................................................................................................................16

3.1 Advantages of Big Data Security Analytics.......................................................................16

3.2 Challenges of Big Data Analytics in retail.........................................................................17

4. 0 Implications.....................................................................................................................................18

4.1 Big Data Analytic : Case Study

4.1.1 Macy’s Dynamic Price Optimization…..................................................................... 18

4.1.2 Tesco’s Inventory Management..………………………………...………………….19

4.1.3 Kroger’s Promotion Analysis……………………………………………………….19

5. 0 Conclusion..................................................................................................................................... 20

References............................................................................................................................................. 21

**List of Figures**

**1.** Figure 1: Retail Application Process……………………………………………………….9

**2.** Figure 3: Agent-Based Big data analytics ………………………………………………...13

**3.** Figure 4: Retailer’s Biggest Obstacle to Success with Analytics ……………………...…14

**4.** Figure 5: Big Data’s Impact on Retail Process………………………………....................14

**5.** Figure 6: Retail functions with gain benefit from big data………………………………..15

**6.** Figure 7: Retail operations using big data analytics……………………………………….18

**ABSTRACT**

The ability to perform retail operations while rapidly developing the information architecture is critical for retailers who want to enhance trader’s insight into customer behaviour and purchases. Data tsunami has become own share for retail industry to wrestle with. Now days, the number of retailers and their diversity are increasing rapidly. As a result, managing the types of data is becoming more complex. Similarly, velocity and volume of data is increasing which makes it difficult to extract the relevant value from big data. To overcome with this problem, that is to obtain valuable insights from their diverse data and grab customers in competitive marketplace, applying predictive and prescriptive data analytics in real time is a perfect solution. Thus, Adoption of big data in retail industry and big data analytics for overcoming retailer’s business challenges is the aim of this paper. Also to deal with these difficulties agent based paradigm is proposed in this paper. This paper delivers the overview about meeting customer needs by adopting big data and analytic capabilities in retail environment. In short, this paper will explain about how retailers view big data and to what extent they use it to benefit their businesses.

**Keywords: Big Data, Big Data Analytics, Retail, Agent-Based Data Analytics**

1. **INTRODUCTION**

“Information is going to be our generation’s next natural resource like steam was to the 19th century” (Terry Lundgren CEO, Macy’s). In recent years, the world is producing more data ever before as the world is becoming digital. Digitalized world produce large amount of data from social networks and internet of things which is complex in structure. Generally, the data generates from different sources like emails, search queries, online transactions, logs, audios, click streams, videos, sensors, mobile phones etc. It has becoming very difficult for industry to store such a massive amount of data and use it accurately for performing the operations on it. The solution for performing the operations like storage, access, pattern recognition and analysis on this data should be invented. As the data is very large, the term use to call this data is “Big Data” and term for operations perform on this data is called as “Big data analytics”.

**1.1 What is Big Data?**

Big data is also defined as large amount of data arriving from multiple sources that need to be handled by new technologies so that by using capturing and analysis process, the expected values can be extracted. (James Manyika, et al. Big data).

**Characteristics of data:**

1. Usually, big data work with Petabytes or Exabyte of data. It also has the potential to gather records in billions and trillions.
2. Some flat schemas and complex inter-relationships are present in big data. Not only is this, but time-stamped events and incomplete data sets present in big data.
3. Some incidental connections between the data sets are also included in big data.

**1.2 Why is big data important?**

Big data contain terabytes, petabytes, or maybe even Exabyte of data.  It is highly difficult to keep track of such data. (GregSatell, 2013). The traditional “scientific” studies and “statistical significance” was relied on before to determine information the companies could trust. Big data handles the data which is stored in traditional warehouses as well as the data not suitable to be stored in those warehouses. Big data does analysis on whole data available which points out the point of failure and increases the longevity of log storage. Big data analytics expose and visualize hidden patterns and also analyse the knowledge produced for decision making. Big data is important as it is important to calculate the risks as a lot of data is under-utilized. The biggest use of big data is the social media and customer sentiments as it is necessary to keep an eye on customer’s viewpoints.

**Types of data in Big Data are:**

**Traditional enterprise data**: It is a data which consist of customer information from customer relationship management systems, Web store transactions, or transactional ERP data.

**Machine-generated or sensor data**: This is a type of data which includes different Call Detail Records (CDR), some smart meters, different types of weblogs, sensors, etc.

**Social data**: It is a data may be generated from different types of blogging websites such as Twitter, or different social media platforms such as Facebook, Orkut etc.(Liu, Yang, & Zhang, 2013)

1. **LITERATURE REVIEW**
   1. **Big Data Used in Retail :**

* The process of separating wheat from chaff can be easy way to define retail big data analytics. (DeZyre, 2015). For retailers, big data is very differentiating and promising. It is like an asset for them only if they make sense of it. To create more specific relation with customer and to coordinate the fulfilment of complicated operations, the right information at right time is more important. Critical data generated from inventory management, service, sales, returns, marketing, finance, distribution, assortment planning etc. must be optimized to become more profitable and cost effective by continuous directions. This type of trend is increasing day-by-day that the prediction of increasing number of consumers which will use internet for everything becomes true by 2020. (IBM Corporation White paper, 2013)
* To grow in market, retailers should collect the valuable data and should ensure that no relevant data is lost. Retailers achieve 73% higher sales than others by using predictive analytics. Developing customer centric capabilities is crucial in today’s retail world where the customer is in the driver’s seat. To grow economically, retailers should understand, predict and deliver what customer needs. This can be done by implementing new technologies and capitalizing new data sources. Thus, it is important for retailers to have 360 degree view of customers. (DeZyre, 2015)
* Now days, Consumers are becoming more digitalized that they use ecommerce platforms to find their product with best offers, recommendations, reviews on their favourite brand etc. Because of this readily available information, consumers can better able to compare the products, services and prices online same as they do in physical store. This gives trigger for increase in the competition of all retailers in market. To be the best in the market, retailers should collect the accurate information about their customers, manage the information and analyse it by analysing velocity, volume and variety of data. It allows companies to focus on generating more revenue that is a positive transformation to serve the business in better way.
  1. **Need for Retail Big Data Analytics :** (DeZyre, 2015)
* Lot of data has been generated at retailer’s side. Millions of people order lot of products from many supermarkets. These millions of people generate millions of data which can be handled by legacy system is very difficult. Apparently, there will be a data loss as the speed of legacy system is very limited. Whereas the use of big data analytics makes it very easy to analyse the data and sort it out easily for retailer’s revenue growth.
* Big data analytics not only make easy analysis for customer but it also makes it easier for retailers to analyse the customer reviews and create a specific offers to increase their sell. Social media makes it easier for consumers to compare the different products and its prices online or in retail stores. Consumers can also interact with the retail channels through social media and can make a shift from one brand to another brand by observing and understanding their needs and online reviews, comments or tweets.
  1. **Retail application process:** (Agent-Based Big Data Analytics in Retailing : Case Study, 2015)
* In the past, retailers used to have supermalls and outlets where they used to attract the customers by lot of advertisements. To attract the customers, they have also designed their stores with lighting and colours to provide comfortable and wonderful environment to customer. For proper analysis, the customer profile should be created and after creation of this, retailers used to perform seasonal sales analytics, market-basket analytics, pricing optimization analytics and inventory optimization analytics.
* The result of this analysis provides insights regarding behaviour and buying patterns of customer. The retailer is always curious about reason behind the lot of shopping, pricing, inventory, spending and clickstream promotion data which they have. They are also curious about their product reviews from the customers.
* Retailing is now moving rapidly from physical in-store model to multi-screen and multi-channel experiences. This change of getting everything online by using internet is giving tremendous challenge to traditional in-store approach.

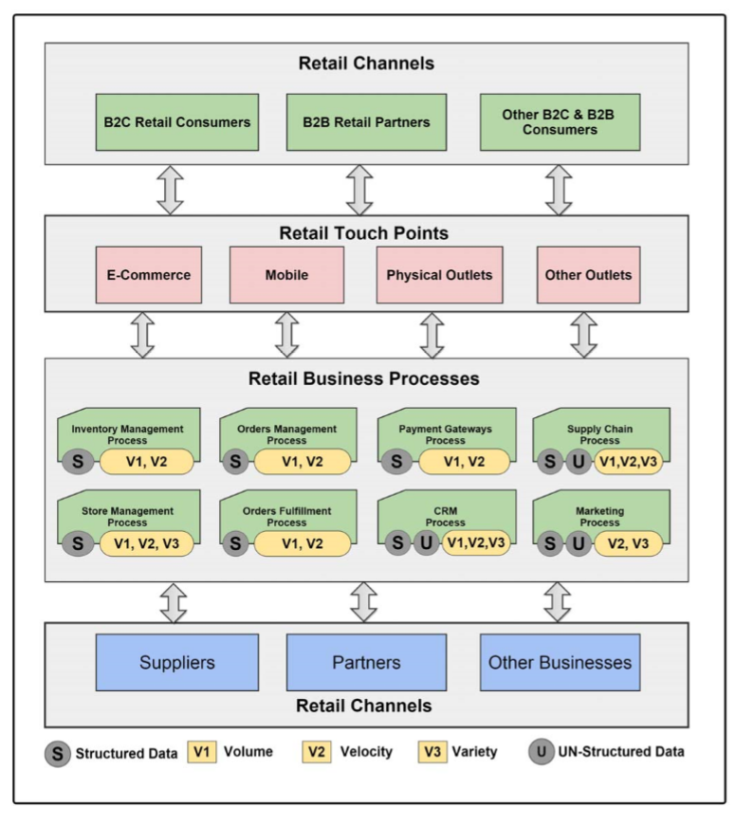


Figure 1: Retail Application Process (Agent-Based Big Data Analytics in Retailing : Case Study, 2015)

* Because of this tremendously and quickly generated data, the retailer got a challenge to analyse it and spot trends hidden within digital clue before their competitor discover them. This is where the multi-channel analytics involved for sophisticated targeting of customer behaviour.
  1. **Working of Big data in retail** 
     1. **Data: -** Big data in the retail field is the information coming from different sources. The data can be

1. A personal information like name, age, date of birth, transaction history, store visit frequency etc.
2. Data coming from social media like tweets, reviews, comments etc.
   * 1. **Big data processing steps:**
3. Big data is generally stored into big data repository.
4. Then this raw data is processed using business analytics software to make it actionable data.
5. This actionable data then can be used for demand forecasting, inventory forecasting, trend forecasting, dynamic pricing, planning product promotion and timely offers.
   * 1. **Working of big data analytics:**
6. It identifies the prospect via smartphones etc.
7. Then it look for prospect location and profile information into repository.
8. Feeds the profile of prospect into real time expert system.
9. In real time, determine best offer for prospect customer.

This is how the analytics work using big data and determine the different analytics for retailers to improve their revenue and growth in market.

* 1. **Five use cases in Retail** (QBurst, 2015)

1. **Personalization :**

* To increase revenue is the main motivation of any retailer. There are lots of ways to fulfil it. In which, one way is at personalize level. There are independent retailers at personalize level who envision data personalization as a most important driver of increased revenue in near future. There are total 57% independent retailers from which 22% already making use of this data personalization where as another 35% are planning to use it as soon as possible.
* For data personalization , there are some analytical techniques like behavioural targeting, purchase pattern analysis, recommendation of products, psychographic segmentation, different individual offers etc. applying online and in store. To publicize personalize content on websites, machine learning techniques and collaborative filtering algorithms are used. This process comes under Recommendation engines.

1. **E-Commerce Optimization :**

* For independent retailers ecommerce is very small platform to increase their revenue. So, optimizing this ecommerce and transpire it into business function create the big opportunity for improve impact of analytic on the revenue.
* To obtain greater engagement of customer with good conversion rates, retailers reform product landing pages by driving into website analytics and clickstream data.

1. **Brand Evaluation :**

* Brand evaluation can be easily done by using social media channels and online networking platforms.
* As mentioned earlier in this paper, customers post their views about the product on social media channels. The delighted or frustrated views, posts , tweets or other social content if analyse carefully by retailers using natural language processing tool will gives the result of acceptance of a brand by customer.
* To analyse these sentiments of customer from their feedback, social sentiment analyser is used by retailers. This analyser categorise the sentiments in 3 categories as positive, negative and neutral. After the marketing strategies done by the retailers, this analyser will be useful to monitor these sentiments in real time from the customer’s feedback and will create a respective solution to improve revenues.

1. **In-Store shopping experience:**

* Even though everything is becoming digitalized, shoppers will like to go physically to the retail shops to not only buy the products, but to socialize and to satisfy their senses. In this case, it is the responsibility of retailer to combine the data from online, social and mobile channels and perform analytics on it and provide the same online solution within the physical store too.
* There are many more ways to let customers get engaged. One way to engage the customer with the same brand is to measure the effectiveness of advertisement and plan in-store marketing campaigns. So, to perform this task information can be taken from multiple touch points, customer service agents or surveillance cameras etc.

1. **Dynamic Pricing**

* Depending on demand and sale of the product the online merchants decide whether to lower or hike the price of the product. This can be used by price optimization software. Stock keeping unit used to change the price manually on which the retail industry used to work till now. But this system will not work today with new technologies. So, the repricing engine came into the market which monitors market trends, consumer demands and competitor prices to generate dynamic prices.
* For the best revenue for retailer according to their market price this repricing engine shows the maximum and minimum limit of price which can be used for setting the price of product and to display it to shoppers.
  1. **Proposed Agent-based big data analytics in retail :** (Agent-Based Big Data Analytics in Retailing : Case Study, 2015)

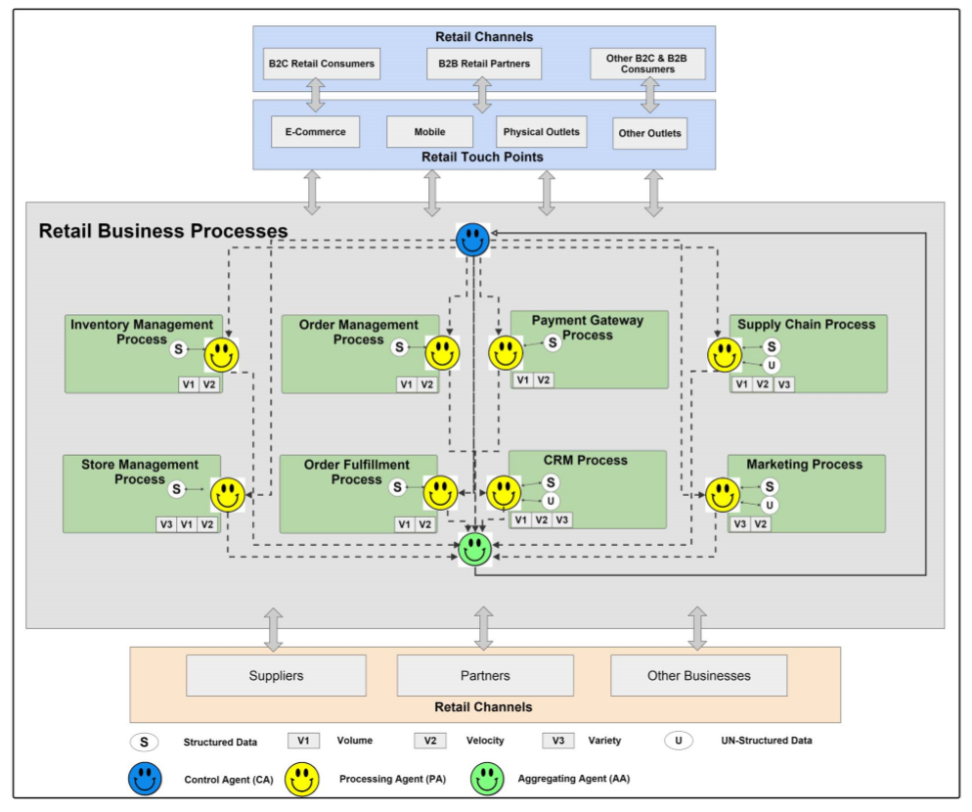


Figure 2: Agent-Based Big Data Analytics

Reduction of unrelated data or the nature of dataset which is not uniform across datasets is very necessary. Also for further analysis and inference the reduction of data is very necessary. So the agent-based big data analytic scheme is used for abstraction in big data. The abstraction is only the reduction of data which includes identify representative patterns, large data sets analytics, reduction of attributes and features , heterogeneous data set access and integration. Because of the heterogeneous nature of data set, hierarchical clustering method is useful for this method.

**Parts of this proposed approach:**

1. **Control Agent(CA) :-** In retailing process, to control the access from multi-channels, CA is nominated. By coordinating with other agents, data sets that flow inside the retailing process are managed.
2. **Processing Agent (PA) :-** The responsibility of PA is to classify the data on the basis of resource type, but filtering dataset from disposable ones and extraction of input datasets.
3. **Aggregating Agent (AA) :-** There must be some agent to gather and integrate all valuable data and knowledge which was extracted by effective method from retailing process which can enable decision making.
   1. **Retailer’s Biggest Obstacles to Success with Analytics** (2014 Big Data in Retail Study, 2014)

Obstacles retailers are getting after dealing with big data analytics made by the executives. They responded as the 41% need is for ‘single version of the truth’ followed by 38% need is for inability to analyse data at a low enough level of detail and 34% need is for difficulty accessing and integrating enterprise. Then Significant challenges includes 16% slow query speeds, 15% of mandatory reporting tools that hardly handle sophistication of retailer’s analytical questions and 13% need for lack of self-service reporting.

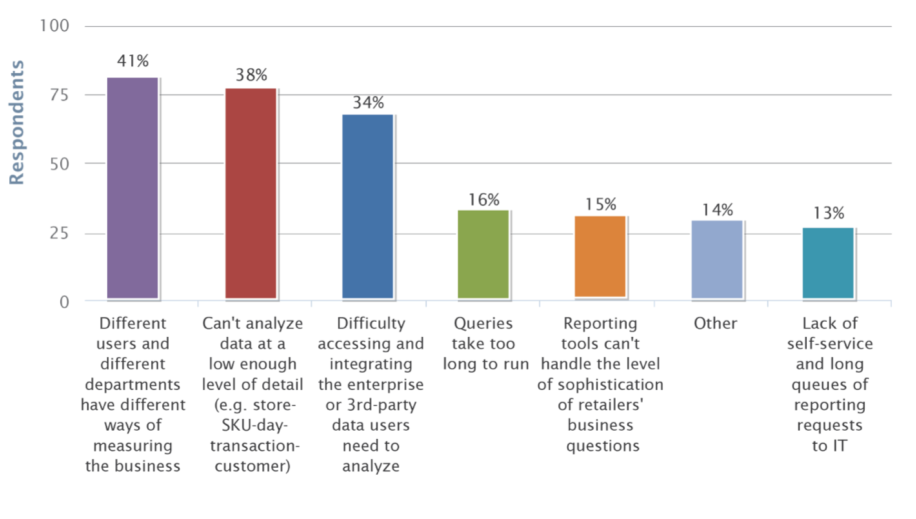


Figure 3: Retailer’s Biggest Obstacle To Success with Analytics

* 1. **Big Data’s Impact on Retail Processes** (2014 Big Data in Retail Study, 2014)

Impact of Big data technology on specific retail processes are as follows. The design of targeted offers and promotions has 50% impact whereas forecasting and supply chain modelling has 49%, customer-centric merchandising has 43% and loyalty program management has 35% impact of big data technology. Retailer also interviewed to analyse the benefits of using big data in percentage gives the result as 28 % in workforce management, 16 % in loss prevention and 18% in store design. However, 5% of retailers felt that impact of big data analytics will not affect retail business processes.

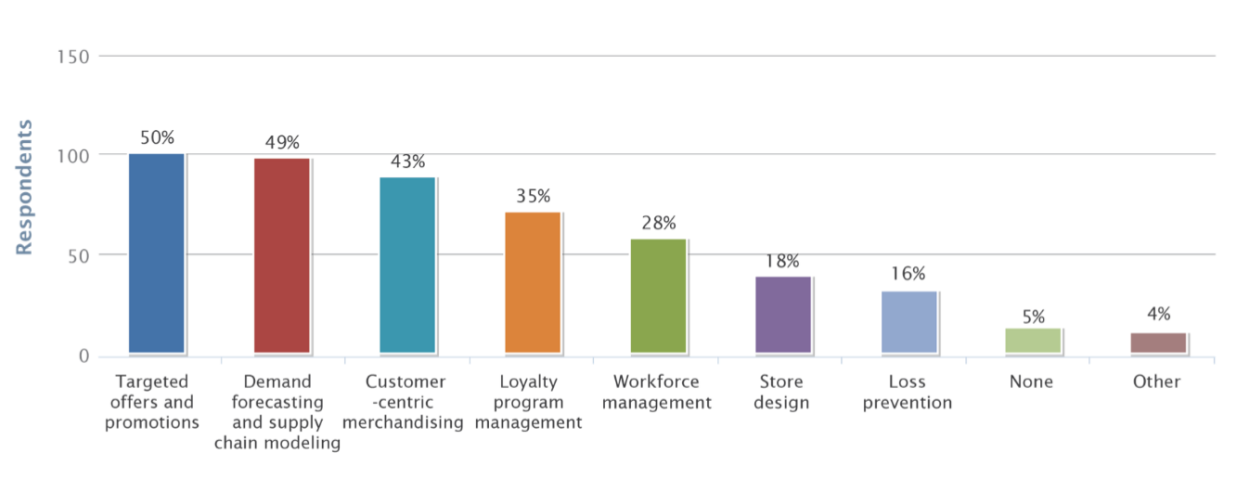


Figure 4: Big Data’s Impact on Retail Process

* 1. **Retail functions with Gain benefit from Big data** (2014 Big Data in Retail Study, 2014)

Above all retail functions, merchandising and marketing come ahead by gaining 53% and 48% big data benefit respectively. Then functions which followed closely are store operation and E-Commerce gains 42% benefits each. Then 27% , 23% and 21% benefits gained by supply chain, finance, loss prevention respectively.

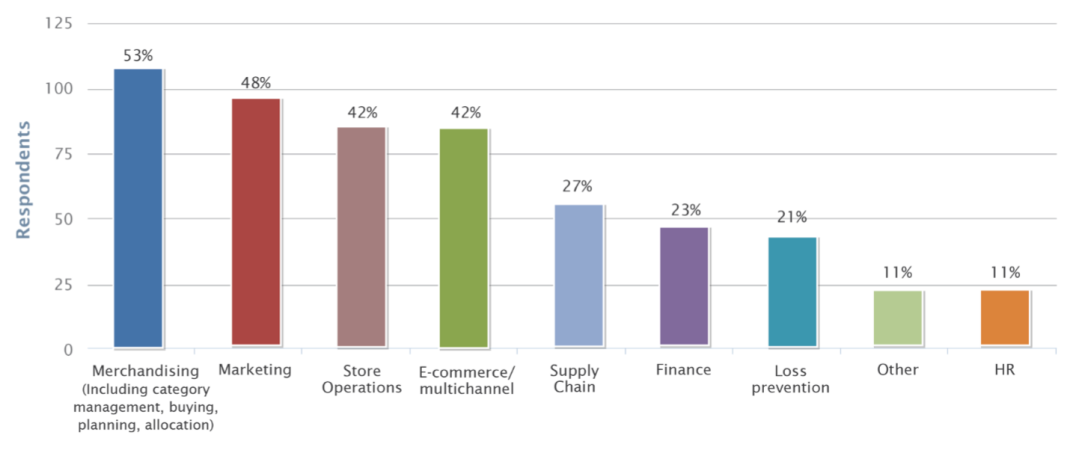


Figure 5 : Retail functions with gain benefit from big data

1. **DISCUSSIONS**
   1. **Advantages of Big data in Retail**
2. **Increases Speed of Analysis**

* The revenue optimization for entire organization which used to take lot of time can be completed now in less than two hours.
* Earlier, it was not feasible to generate a daily forecast and facilitate immediate decisions for executives but now it has been very easy for them.

1. **More Strategic and Combine work of IT and Business**

* For every need of technology, Business users forced to take advice of IT people. Retail IT organizations facing large amount of new needs and technologies. But now there is High-Performance retailing which helps both. Separately IT retailing can create retail analytics foundation whereas Business users can help merchants by visualizing, hypothesizing and making a model from innovative ideas using science.

1. **Give real-time decision-making to retailers**

* Customers when walk into stores expect best service from store managers and associates. To achieve this game-changing decisions and actions are required which can be enabled by high performance retailing only.

1. **Avoid offer spam**

* Till now, most of the customers faced a problem of getting many messages having lots of offers even when they don’t responds to it. It may harm the brand reputation in market. To avoid that, retailers are using high performance computing due to which companies can only send these messages to the target shoppers who previously responded to the same.
  1. **Challenges of Big Data Analytics in Retail**
* **Customer Predictions**: It is very important for retailers to predict the customer’s demands about the product as well as predicting the pattern of what customer will likely to buy with other products. However, as the numbers of customers are increasing, it has become difficult for the retailer to predict their customer’s buying habits.
* **Changing Trends :** Once the retailer analyse the pattern of customer , it is easy to apply that and increase the revenue from it but recognising pattern and applying should be as fast as possible because the trends are changing very frequently. Trends are changing so quickly because of communications, changing technologies and different consumer tastes. It is becoming very difficult for the retailers to analyse the patterns very quickly as trend changes.
* **Sentiment Analysis for Market Reputation:** We usually see that rumours spread very quickly than the compliments. Same in the case of customers and retailers that if customer is unhappy with the product then product will not run in the market whereas if the customer is happy then it will take some time but product will definitely get success and retailer will definitely enhance their growth in market. Customers never forgive for these mistakes. So to make accurate and precise prediction for happiness of customer, retailors have to do sentiment analysis using Hadoop.

1. **IMPLICATIONS**
   1. **Big Data Analytic: Case Study** (Delhaize, 2016)

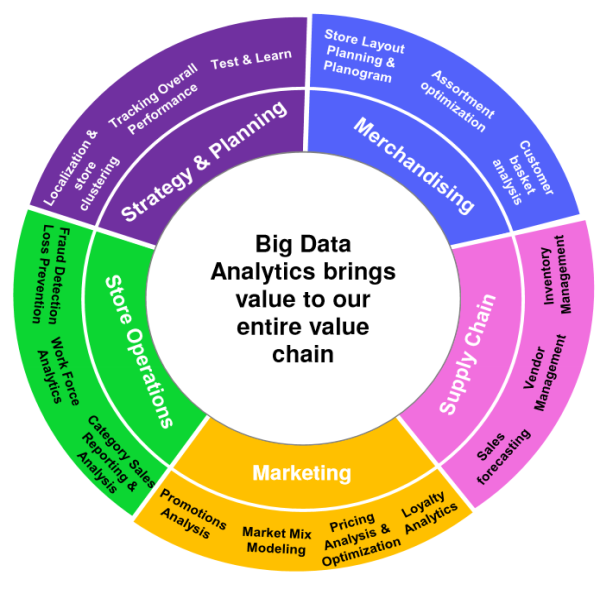


Figure 6: Retail operations using big data analytics

For some operations mentioned in figure, there are some real-time examples:

* + 1. **Macy’s Dynamic Price Optimization :**

Macy’s store is the famous store in the united states of America. It has 876 stores in 45 states. To determine optimal pricing, more than 2 terrabytes of price information analyses by macy’s. Analysing speed is also increased more than 60-fold as it moved to cloud-based solution. Macy’s store auto-adjust pricing for higher sale. On daily basis, it analyses 270 million different factors in 20 minutes across 73 million products. The time for analysis here is now reduced from 22 hours to 20 minutes.

* + 1. **Tesco’s Inventory Management:**

Tesco is a store where when they approach sell-by dates reduces the price of half a billion products every year. The products which are more price elastic than others can have later price reducing than others. Tesco made some models which they runs to see whether how much price they should drop and when they should shift the product and reduce its waste. In case of Stock Management, Tesco in average distribution centers has capacity to hold 12000 items from 700 suppliers. Traditional policy is that if you order in bulk then you will get lot of discount. But using Big data analytics in Matlab which runs full simulation Tesco tackle with individual pieces.

* + 1. **Kroger’s Promotion Analysis:**

As a CEO of Kroger, big data is a secret weapon in grocery store is proved by David Dillon. 120 analysts work full time n Kroger business. From 40 billion purchases made from 4 billion shopping trips 300 terabytes of data sifts by teams over 2 years. Annual revenue is in excess of $100 million managed only because of big data analytics. Kroger has an reputation in the market of having most sophisticated and influential retailer databases in US. Mailing of coupons to 11 million households on regular basis from which 97% are personalized has done by Kroger. Kroger has different focus of growing basket size than acquiring new individual customers.

1. **CONCLUSION**

I have outlined the big data strategies with its use in retail. I have also mentioned the need of big data in retail companies. To improve the growth in the market with increased revenue, the retailer started using big data analytics. I have also described the challenges and benefits of big data in retail. Big data approaches developed by IT companies can also apply in retail domain. I have also specified the technology that may need to be investigated further: the Agent-based big data analytic in retail. The big data usage helped analytic processes to improve their performance and completing the task in very less time. It also helped in decision making and conversion of decisions into action. I have also mentioned different analysis made after interviewing the executives of retail companies. After viewing all the research of the big data in retail, we can reach to the conclusion that the retailers are getting more benefit of the big data technology than challenges they have to face.

# References

1010data. (2014). *2014 Big Data in Retail Study.* Retrieved from www.1010data.com

Datastax Corporation White paper. (2013). Big Data: Beyond the hype. *WP-Datastax-Bigdata*, 4.

Delhaize. (2016, September 25). *Big Data: Impact and applications in Grocery Retail.* Retrieved from Big data - Frans Mullar, Erim : http://www.erim.eur.nl/fileadmin/centre\_content/marketing\_of\_innovation/Big\_Data\_-\_Frans\_Muller.pdf

DeZyre. (2015, April 2). *5 Big Data and Hadoop Use cases in Retail Analytics.* Retrieved from Dezyre Website: https://www.dezyre.com/article/5-big-data-and-hadoop-use-cases-in-retail-analytics/91

Firas D. Ahmed, M. B. (2015). Agent-Based Big Data Analytics in Retailing : Case Study. *IEEE*.

GregSatell. (2013). *Forbes*. Retrieved from http://www.forbes.com/sites/gregsatell/2013/10/11/why-big-data-matters/: http://www.forbes.com/sites/gregsatell/2013/10/11/why-big-data-matters/

IBM Corporation White paper. (2013). Analytics : real-world use of big data in retail. 1-2.

Liu, Yang, & Zhang. (2013). A Sketch of Big Data Technologies. *A Sketch of Big Data Technologies*.

QBurst. (2015). *Application of Big Data in Retail.* Retrieved from htttp://www.qburst.com