

Sentiment Analysis through Big Data in online Retail Industry: A Conceptual Quantitative Study on linkage of Big-Data and Assortment Proactive of Online Retailers

Muhammad Faisal Sultan¹

Mehwish Jabeen²

Muhammad Adeel Mannan³

Abstract

Big-Data is the recent trend in data sciences prevailing all over the globe. The tool aids significantly in optimization of knowledge and has predominant use in optimization of knowledge and productivity. However, there is lack of understanding of concept and its application in Pakistan as indicated by Gallup Pakistan (2018) and stream of data is going to be doubled in two years' time Tankard (2012). Therefore, there is a definite need of research which optimizes understanding associated with technology and its application from the context of Pakistan. Hence considering the application of big-data in retail sector this study aims to explore the impact of sentiment analysis through relating impact of big-data with effective assortment s of online stores. Although data has been collected from IT experts associated with online retail sector via quota sampling and SMART-PLS has been incorporated for the purpose of analysis. Results of the study highlights that big-data is perceived as the major tool for the betterment of assortment in online retail stores although data scientist and their applicability might diminish the impact of the use of big-data.

Keyword: Big-Data, Sentiment Analysis, SMART-PLS, Assortment and Sentiment Analysis.

1 Introduction

Big-Data is not a new terminology which is developed in 2011 after the continuous and thorough efforts made in the field of data management (Surbakti, Wang, Indulska & Sadiq, 2020) In recent times organizations are gathering heaps of data in order to gain benefit in upcoming time. Therefore, there is a necessity of managing this enormous amount of data effectively so to extract optimal information at the time of need (Fan & Bifet, 2013). Although in the time of need one of the latest IT trend trends i.e. Big-Data is achieving massive attention from researchers & practitioners (Bollen, Mao & Zeng, 2011). There are numerous advantages of using big-data which will provide access to new markets and aids in innovation of business model. Moreover, the technology will also aid in assessing customers' needs which also results in betterment of customer service (Rajeb, Rajeb & Keogh, 2020).

2 Statement of Problem

Rajeb (2020) postulated that there are some legitimate indications regarding increase of big-data application in terms of marketing. However, study uses the reference of Pantano Giglio and Keogh (2019) that regard less of increase in literature on big-data on various marketing

¹Khadim Ali Shah Bukhari Institute of Technology (KASBIT) | mfaisal@kasbit.edu.pk

²Khadim Ali Shah Bukhari Institute of Technology (KASBIT) | mehwish.jabeen@kasbit.edu.pk

³Hamdard University | adeel.mannan@hamdard.edu.pk

functions although literature provide minimal evidence for association of big-data & marketing functions. Although it's a time when each and every industry is striving to know how big-data might be associated with the solving problems. In fact, some of the industries have already implemented this technology (Le & Liaw, 2017). However, there are significant lacking in the use of big-data technology in Pakistan and only NADRA has some advanced mechanism of big-data (Ashraf, 2013).

This actually happens due to lack of knowledge and understanding regarding the technology (Gallop Pakistan, 2018) and difference of culture as compared to the western world (Latif, Tunio, Pathan, Jianqiu, Ximei & Sadozai, 2018). Hence there is a significant need of research which may analyze the effect of big-data technology with reference to the business operating in. Pakistan. Especially in the context of marketing as big-data has the ability to transform marketing related functions in near future (Rajeb et al., 2020).

3 Theoretical Framework and Delimitations

Application of big data are linked with top ten industries including wholesale and retail (Rajeb et al., 2020). Among these retailing and wholesaling are treated as part of daily life (Le & Liaw, 2017). On other hand use of information technology might significantly aids in the profitability of retail sector (Ali Subzwari and Tariq, 2016) though there was lack of evidence regarding the application of big-data in marketing functions but have the ability to transform market and interrelated functions in near future (Rajeb et al., 2020).

Hence study will figure out the impact of big-data on marketing functions of retail sector which is associated with exponential growth (Fazl-e-Haider, 2018), and major elements are pricing & assortment (Aktas & Meng, 2017). However, to deal with big-data analytics there is a requirement of technical mix of creativity and analytical skills especially in the context of marketing functions (Glass & Callahan, 2014). Though this is not possible without conducting quantitative analysis to forecast these measures in more effective manner (Aktas & Meng, 2017). Therefore, this study will use availability of skilled data scientist as the moderator with assortment of retail sector of Pakistan.

The research has only one IV (big-data) and one DV (assortment strategies) due to its conduction in Pakistan where there is lacking of understanding of the concept of big-data (Gallup Pakistan, 2018). However, most of the retailers are dealing in fast moving consumer goods (IBM, 2018). and online retailing is still infancy and in initial stages (Ali et a., 2016). Therefore, in order to provide significant impact to economy it is valid to measure it efficiency in terms of online retail stores.

4 Research Questions:

RQ1: Whether Big-Data is applicable to developing countries like Pakistan?

RQ2: Whether knowledge of IT experts really signifies the use of Big-Data?

RQ3: Whether Big-Data is applicable to retail sector of Pakistan?

5 Significance

Rajeb et al. (2020) indicated that use of big-data analytics is viable in top ten industries all over the globe. Although there is lack of understanding regarding the concept and application in the context of Pakistan (Gallop Pakistan, 2018). However other countries are dominant in taking advantage of big-data application which are even observable in leading industries like banking, securities, media, whole selling and retaining etc (Aktas & Meng, 2017).

However whole selling and retailing are treated as the part of daily life Li and Liaw (2017) and therefore conducting quantitative study on implication of big-data on retailing industry must be termed pervasive in nature. The claim is valid as study will fulfill need of big-data linkage with marketing functions as highlighted by Rajeb et al (2020).

Moreover, study will also help in removing the lacking of studies linked with big-data from emerging markets (New Desk, 2020), hence the study is beneficial for academia as well a pragmatic world and must be termed as pervasive.

6 Literature Review

One of the latest studies of 2020 indicated that still a significant lacking of studies which highlights the most appropriate use of big data technology (Surbakti et al., 2020). The terminology is actually associated with large data sets i.e. of terabyte and exabyte (Rajeb, Rajeb & Keogh, 2020). However, there is also a significant probability for the increase of organizational data (Zanini & Dhawan, 2015) thus several studies which tries to explore challenges pertaining to big-data and most of these indicated further investigation of benefits which organizations may urge through using big-data (Surbakti et al., 2020). There was lack of evidence for the relationship of big-data and marketing functions but big-data provide companies way to create edge and value (Zeng & Glaister, 2018)

However recent work of Rajeb et al (2020) highlighted that big-data is playing major role in improvement of functions like Marketing, supply chain as well as in process of decision making. This form of data is also applicable social media to observe thoughts, views and reviews of clients and this field is termed as sentiment analysis with purpose to device effective marketing strategies (Zanini & Dhawan, 2015). Though as mentioned earlier there is lack of evidence for the relationship of big-data and marketing functions (Keogh, 2019) but big-data has significant importance in retail business (Bradlow, Gangwar, Kopalle & Voleti, 2017).

Therefore, the study is optimal to relate study with sentiment analysis as a form of big-data analytics so to collect, organize and analyze large data sets in order to uncover new patterns and large data sets (Vijayarani & Sharmila, 2016)

7 Big-Data & Assortment

Product attributes as well as level of product attributes are the major components of product information at retail stores (Bradlow et al., 2017). Products attributes are potent predictor of customer's satisfaction as through effective analysis firm may decrease the percentage of stock

out products (Matsa, 2011) & may also increase total number of items (Briesch, Chintagunta & Fox, 2009). On the other side retailer are in need of right data to gain information about customer insights and value might be gauged through testing the data in terms of transaction log, loyalty information, pricing strategies and campaign results (Howe, 2014). Analysis of purchase records by Gielens (2014) was based on six retail stores for thousands of customers and resulted in effective negotiation with suppliers.

Similarly, Kumar and Kapoor (2014) indicated buying behavior has been measures on the bases of three characteristics i.e. quantity of purchase, frequency of purchase & preferred location of retail outlet. Thus, these types of information can be used by retailers in order to optimize practices of inventory management and may also deal effectively with change in demands (Ridge, Johnston & O'Donovan, 2016). Big-Data is also capable of assisting retailers in recognition of anomalies through observing practices and patterns looks unusual (Kaur & Jagdev, 2017).

Although several retailers are failed to incorporate effective data collection techniques due to privacy issues but these types of issues must not be treated as reason for non-incorporation of better data driven decisions. On the other hand, if retailers became able to cope up with the opportunity of associating data with business analytics might be able to track entire purchase journey (Shankar, 2019).

Though this might only be possible if data scientists have sufficient skill inventory to deal with issues of knowledge extraction (Dolezel & McLeod, 2019). Hence, there is severe lacking of inhouse data management specialists due to uncertainty of return on investment from big-data analytics (Iqbal, Kazmi, Manzoor, Soomrani, Butt & Shaikh, 2018). In large firms the functions is carried out through using experts of different IT fields. However, small and medium sized enterprises (SME) faces lot of difficulties in managing the same as they need cross sectional experts. Contrary to this unavailability of experts of the field & high cost of staffing are also resulting in the shortage of in-house data management specialists (Iqbal et al., 2018).

8 Research Methodology

Research methodology is a generic logic used in devising research although methods are the specific set of strategies and procedure used in the process of analysis. Research methodology is also supplemented with epistemological or ontological assumptions (Long, 2014). However, if we use facts as a truth then epistemology would help us in posing factual questions i.e. how do we know the truth? and what counts as knowledge? etc.

Although ontology is a branch of philosophy through which we evaluate things which making sense are real or not and through this researcher tries to explore (Kivunja & Kuyini, 2017). Hence in association with parameters of Kivunja and Kuyini (2017) the philosophy associated with this study is epistemology. Reason being the study aims to identify linkage of big-data in terms of retail industry of Pakistan in order to seek knowledge rather than to check the nature of reality. For discussing further on research methodology, one has to discuss about two of its parts i.e. Research Design & Sampling Design as indicated by Sileyew (2019).

A Research Design

Research design the part of research methodology which is used to provide answer to those questions in which researchers are interested (Oso & Onen, 2009). On the other hand, Mkansi and Acheampong (2012) indicated that most of the research terminologies which are used in academic research are consistent with the work done by Saunders Lewis and Thornhill (2009). Therefore, most of the terminologies used in research design are consistent with Saunders et al., (2009). Therefore, the philosophical stance is required for the determination of most adequate method for collection and analysis (Zukauskas, Vveinhardt & Andriukaiteiene, 2018) which is applicable to qualitative as well as quantitative research design (Saunders, Lewis & Thornhill 2015). Though the method of data collection is mono-method i.e. quantities (Saunders et al., 2015) and unit of analysis is individual while time horizon was cross-sectional (Sekaran & Bougie, 2016).

B Sampling Design

The reason due to which one may realize the base to include specific units' items in research are termed as sampling design (Mugenda, 20003).

However, in order to decrease overall cost associated with data collection items which are included in sample must able to assure research objectives (Leedy & Ormrod, 2005). Therefore, this study uses IT experts from online retail sector as the elements of sampling regardless of the infancy of online retailing (Ali et al., 2016). However, most of the retailers in Pakistan do not preferred online method (IBM, 2018) hence there are few IT experts available who are associated with online retail business. Therefore, the sample size for this study is 50 respondents which are justifiable as the research is linked with theory building approach due to its linkage with big data application on online retail sector. However, the data has been collected through using quota sampling so to deal with slow response rate and also to excessive sampling (Yang & Banamah, 2014).

C Questionnaire

Data has been done through of closed ended questionnaire which is adapted from Le and Liaw (2017) and Seetharaman Niranjana Tandon and Saravanan (2016). Moreover, measures indicated by Aktas and Meng (2017) and Valchanov (2017) has also been added to the questionnaire. Last but not the least availability of data scientist is a form of construct which has been developed on the bases of characteristics of data scientist and their use indicated by De Mauro Greco Grimaldi and Nobili (2016). Thus, through these considering systematic pattern indicated by prior studies the questionnaire was formulated.

D Statistical Testing and Analysis

Now a days SMART-PLS is treated as the better option for statistical testing especially in the studies of management sciences (Benitez, Henseler, Castillo & Schuberth, 2020).

Software use two types of models i.e. reflective and formative for the analysis of data (Benitez, Henseler, Castillo & Schuberth, 2020). The model of this study is a form of reflective-measurement model and therefore analysis will follow the indications made by Afthanorhan (2014) and Benitez et al. (2020). Hence the initial tables included in the section will demonstrate about descriptive statistical measures and the later one will indicate inferential statistical measures.

E Outer Loadings

Table 1: Outer Loading

	Assortment	Big-Data	Moderating Effect 1	Skilled Data Scientists
A1	0.743			
A2	0.893			
A3	0.831			
A4	0.891			
A5	0.704			
BD1		0.866		
BD2		0.877		
BD3		0.934		
BD4		0.939		
BD5		0.918		
Big-Data * Skilled Data Scientists			1.166	
SD1				0.686
SD2				0.879
SD3				0.782
SD4				0.876
SD5				0.721

The purpose of table 1 is to indicate outer loading for each element in order to highlight reliability of the construct associated with big-data and assortment strategies of online retailers. The minimum value which may validate the selection of any element in the construct is 0.60 (Afthanorhan, 2014).

Though the optimal range of value for outer loading starts from 0.708 (Sarstedt, Ringle & Mena, 2012) and if the value of element is lesser than 0.708 and also causing in decrease of overall reliability of construct then it must be deleted (Hair Jr, Hult, Ringle & Sarstedt, 2016). Hence all the elements include in table 1 seems to be effective as the minimal value of any element included in table 1 is 0.686 which is higher than the conidian indicated by Afthanorhan (2014).

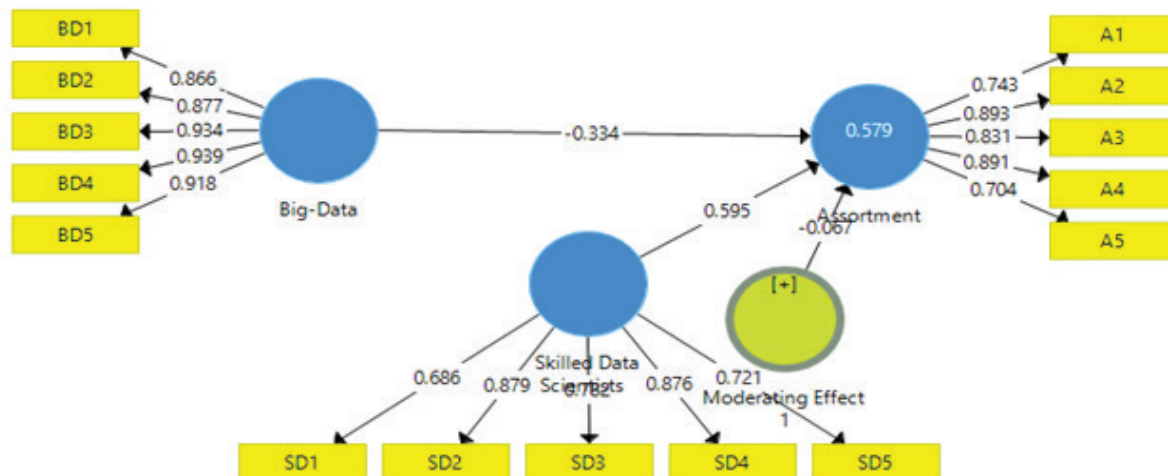


Figure 1: Highlighting p-values to show confirmatory factor analysis (CFA) for the model of big data on assortment strategies

Table 2: Predictive Accuracy (Quality Criteria)

R Square		
	R Square	R Square Adjusted
Assortment	0.579	0.574

Table 2 is termed as quality criteria or predictive accuracy and the use of the tool is to indicate the degree of explained variance caused by independent variable (Benitez et al., 2020).

However the method of evaluation of R is same as that of regression (Andreev, Heart, Moaz & Pliskin, 2009) the minimal value for the tool is 0.26 & of 0.5 & 0.75 are treated as moderate and extensive (Cheah, Memon, Chuah, Ting & Ramayah, 2018). Therefore in accordance with these measure the value of R is treated as moderate as the value for the tool is 0.574 which is lower than the standard of extensive fit.

F Construct Reliability and Validity

Table 3: Construct Reliability & Convergent Validity

	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted (AVE)
Assortment	0.871	0.873	0.908	0.666
Big-Data	0.946	0.950	0.959	0.823
Moderating Effect 1	1.000	1.000	1.000	1.000
Skilled Data Scientists	0.849	0.853	0.893	0.628

Table 3 is indicating construct reliability & convergent validity (Ab Hamid, Sami & Sidek, 2017 & Sijtsma, 2009 a&b).

Table 3 is used to indicate convergent validity that indicates how well parameters associated with one latent variable measures the same construct (Benitez, Henseler, Castillo & Schubert, 2020). Though table also contains two reliability measures i.e. Cronbach's Alpha (α) and Dillon-Goldstein rho in order to highlight construct reliability through Cronbach's Alpha (α), Dillon-Goldstein's rho & AVE (Sijtsma, 2009a&b). Therefore, in the light of these measures the model is effective to ensure construct reliability and convergent validity. These statements are valid as the values of α and Dillon-Goldstein rho is more than 0.7 to ensure construct reliability while values of composite reliability and AVE are more than 0.5.

Table 4: Discriminant Validity through Heterotrait-Monotrait Ratio (HTMT)

	Assortment	Big-Data	Moderating Effect 1	Skilled Data Scientists
Assortment				
Big-Data	0.550			
Moderating Effect 1	0.078	0.133		
Skilled Data Scientists	0.779	0.288	0.140	

Table 4 is used to indicate discriminant validity to show heterogeneousness among different variable of same research model (Cheung & Lee, 2010). Reason for using HTMT as the way to show discriminant validity is the ratio is treated as most effective for highlighting that (Benitez et al., 2020). However, the peak value which is tolerate able in the case of HTMT is 0.85 (Hair Jr, Sarstedt, Ringle & Gudergan, 2017) and any value greater than this is ineffective for highlighting the discriminant validity.

G Mean, STDEV, T-Values, P-Values

Table 5: Path Coefficient

	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics (O/STDEV)	P Values
Big-Data -> Assortment	-0.334	-0.334	0.053	6.290	0.000
Moderating Effect 1 -> Assortment	-0.067	-0.066	0.033	2.062	0.040
Skilled Data Scientists -> Assortment	0.595	0.594	0.054	11.114	0.000

Table 5 is the table which is used to highlight inferential measures of big-data on the assortment strategies of e-retailers. Inferential stats is one of the major part of the measurement models used in SMART-PLS (Hair, Risher, Sarstedt & Ringle 2019). Software does this through t-statistics (Durate & Amaro, 2018) and p-values (Kock & Hadaya, 2018) in order to indicate relationship of variables inferentially.

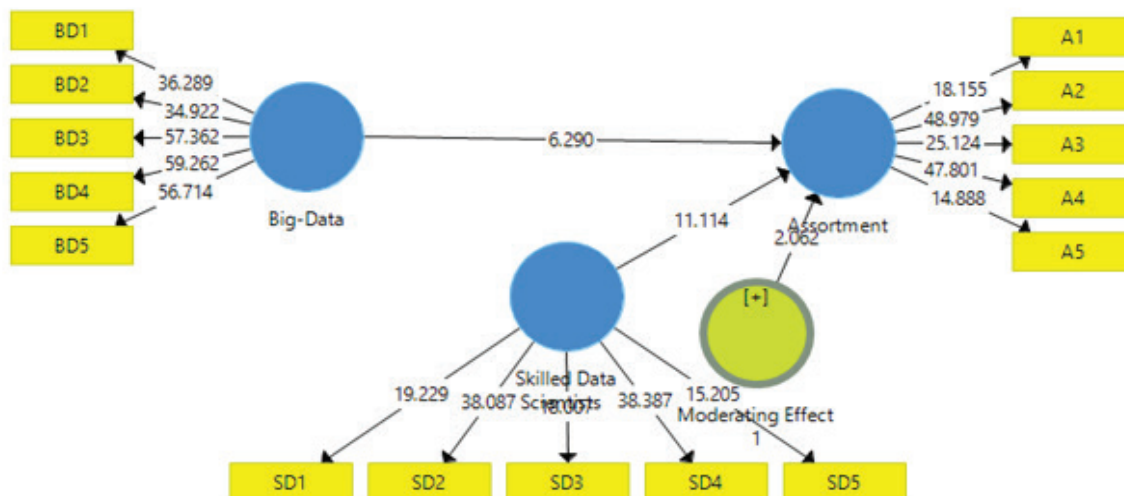


Figure 2: Path Coefficients and regressions weights for the construct of Big-Data on Assortment Strategies of Online Retailers

The benchmark value for t-statistics is 1.97 (Hair, Ringle & Sarstedt (2011) and for p-values the cutoff value is 0.05 above which there is no effect of variable on another (Kock & Hadaya, 2018). The minimum t-value required to indicate relationship between variables of the construct is 1.97 (Hair et al., 2011). Therefore in the light of these parameters it is legitimate to indicate that big-data is perceived as an effective tool for the optimization of assortment of e-retailers and skills of data scientist are also perceived as crucial in optimization of assortment practices.

Although the abilities of data scientist working with e-retailers are not adequate to affect assortment practices at e-retailers. Thus the moderation of skilled data scientist are diminishing the effect of big-dat technology on assortment practices of e-retailres which can be I observed through the increase of p-value in case of moderation of skilled data scientist

9 Discussion And Managerial Implications:

Study highlighted that the indications made by prior studies like Rajeb et al (2020) are appropriate to highlight significant impact of bi-data on betterment of functions like marketing, & supply chain etc. Moreover the study is based on e-retailers thus positive relation of big-data on assortment also indicates that big-data is also applicable to social media. Hence optimal to observe thoughts views and reviews of clients as indicated by Zanini & Dhawan (2015). Thus also justifies the indication of Bradlow wt al. (2017) that big-data is perceived as important tool for the retail business & findings of the study opposes the Keogh (2019). This is legitimate as assortment is important tool for retailers as well as in the field of marketing.

Hence legitimate to declare big-data is a potent tool for the analysis of large data sets (Vijayarani et al., 2016) which also required in retail business to assess right information so to predict customers. Furthermore the moderation of skilled data scientist actually diminishes the impact of big-data from assortment strategies of online retailers therefore also linked with Shankar et al. (2019) and Iqbal et al (2018).

Hence legitimate to declare the ineffective skill inventory of data scientist might be due to imbalance of payments and returns on big-data technology.

10 Area For Future Research

This research is based on the analysis of perceived uses of big-data and results achieved are only from the IT managers and experts associated with online retail sector. However most of the retailers preferred bricks and mortars as business model therefore studies would become legitimate if able to relate big-data with the operations and practices of physical retail businesses. Moreover future studies might try to explore the difference in big-data practices of those retailers which prefer bricks and clicks and flips and clicks in order to check the application of big-data in hybrid form and online form of retailing.

References

- [1] Ab Hamid, M. R., Sami, W., & Sidek, M. M. (2017, September). Discriminant validity assessment: Use of Fornell & Larcker criterion versus HTMT criterion. In *Journal of Physics: Conference Series* (Vol. 890, No. 1, p. 012163). IOP Publishing
- [2] Afthanorhan, W. M. A. B. W. (2014). Hierarchical component using reflective-formative measurement model in partial least square structural equation modeling (PLS-Sem). *International Journal of Mathematics*, 2(2), 33-49
- [3] Aktas, E., & Meng, Y. (2017). An Exploration of Big Data Practices in Retail Sector. *Logistics*, 1(2), 12
- [4] Ali, R., Subzwari, M., & Tariq, S. (2016). Impact of Information Technology on Retail Sector in Pakistan. *KASBIT Journal of Management & Social Science*, 9(1), 63-93.
- [5] Andreev, P., Heart, T., Maoz, H., & Pliskin, N. (2009). Validating formative partial least squares (PLS) models: methodological review and empirical illustration. *ICIS 2009 proceedings*, 193
- [6] Ashraf, S., (2013), "Can we Optimize Pakistan via Big Data?", *Tech Juice*, <https://www.techjuice.pk/can-optimize-pakistan-via-big-data/>
- [7] Baker, W., Kiewell, D., & Winkler, G. (2014). Using big data to make better pricing decisions. *McKinsey Analysis*. McKinsey & Company, <https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/using-big-data-to-make-better-pricing-decisions#>
- [8] Benitez, J., Henseler, J., Castillo, A., & Schubert, F. (2020). How to perform and report an impactful analysis using partial least squares: Guidelines for confirmatory and explanatory IS research. *Information & Management*, 57(2), 103168

- [9] Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of computational science*, 2(1), 1-8.
- [10] Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*, 15(5), 662-679
- [11] Bradlow, E. T., Gangwar, M., Kopalle, P., & Voleti, S. (2017). The role of big data and predictive analytics in retailing. *Journal of Retailing*, 93(1), 79-95
- [12] Briesch, R. A., Chintagunta, P. K., & Fox, E. J. (2009). How does assortment affect grocery store choice? *Journal of Marketing research*, 46(2), 176-189
- [13] Brown, B. Bughin, J., Byers, A.H.; Chui, M.; Dobbs, R.; Manyika, J. and Roxburgh, C., (2011), "Big Data: The Next Frontier for Innovation, Competition, and Productivity", Technical Report for Mckinsey& Company, Washington, DC, USA
- [14] Cheah, J. H., Memon, M. A., Chuah, F., Ting, H., & Ramayah, T. (2018). Assessing reflective models in marketing research: A comparison between PLS and PLSC estimates. *International Journal of Business and Society*, 19(1), 139-163
- [15] Chong, A. Y. L., Ch'ng, E., Liu, M. J., & Li, B. (2017). Predicting consumer product demands via BigData: the roles of online promotional marketing and online reviews. *International Journal of Production Research*, 55(17), 5142-5156, doi: 10.1080/00207543.2015.1066519
- [16] Clark, R., & Vincent, N. (2012). Capacity-contingent pricing and competition in the airline industry. *Journal of Air Transport Management*, 24, 7-11
- [17] Cohen, L., Manion, L., & Morrison, K. (2007). Observation. *Research methods in education*, 6, 396-412
- [18] Czaja, S. J., Charness, N., Fisk, A. D., Hertzog, C., Nair, S. N., Rogers, W. A., & Sharit, J. (2006). Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and aging*, 21(2), 333.
- [19] Danah, B., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662-679.
- [20] De Mauro, A., Greco, M., Grimaldi, M., & Nobili, G. (2016). Beyond data scientists: a review of big data skills and job families. *Proceedings of IFKAD*, 1844-1857
- [21] Ducange, P., Pecori, R., & Mezzina, P. (2018). A glimpse on big data analytics in the framework of marketing strategies. *Soft Computing*, 22(1), 325-342.
- [22] Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17(3), 37.
- [23] Fazl-e-Haider, S. (2018 February 19). Booming Growth in Pakistan's Retail Sector. *Economist*, <http://www.pakistaneconomist.com/2018/02/19/booming-growth-pakistans-retail-sector/>

- [24] Friedrich, O. Stoler, P. Moritz, M. & Nash, J. N. (1983). Machine of the year: The computer moves. *Time Magazine*, 121(1). 14–28
- [23] Gallup Pakistan, (2018), “Big Data Analysis Reports”. Retrieved from <http://gallup.com.pk/polls/gallup-history-project/big-data-analysis/>
- [24] Gielens, K., Gijsbrechts, E., & Dekimpe, M. G. (2014). Gains and losses of exclusivity in grocery retailing. *International Journal of Research in Marketing*, 31(3), 239-252
- [25] Glass, R., & Callahan, S. (2014). *The Big Data-driven business: How to use big data to win customers, beat competitors, and boost profits*. John Wiley & Sons.
- [26] Grewal, D., & Levy, M. (2007). Retailing research: Past, present, and future. *Journal of retailing*, 83(4), 447-464.
- [27] Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. *Handbook of qualitative research*, 2(163-194), 105.
- [28] Hair Jr, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage publications
- [29] Hair Jr, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2017). *Advanced issues in partial least squares structural equation modeling*. sage publications
- [30] Hair, J. F., Ringle, C. M., & Sarstedt, M. (2013). Partial least squares structural equation modeling: Rigorous applications, better results and higher acceptance. *Long range planning*, 46(1-2), 1-12
- [31] Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2-24
- [32] Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the academy of marketing science*, 40(3), 414-433
- [33] Hair, J.F., Ringle, C.M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *The Journal of Marketing Theory and Practice*, 19(2), 139-152
- [34] Hajirahimova, M. S., & Aliyeva, A. S. (2017a). About Big Data Measurement Methodologies and Indicators. *International Journal of Modern Education and Computer Science*, 9(10), 1
- [35] Hajirahimova, M. S., & Aliyeva, A. S. (2017b). Big Data Initiatives to Developed Countries. *Problems of information society*, 1, 10-19, doi: 10.25045/jpis.v08.i1.02
- [36] Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. In *New challenges to international marketing*. Emerald Group Publishing Limited.
- [37] Howe, K. (2014). Beyond big data: How next-generation shopper analytics and the internet of everything transform the retail business. *Cisco*, 1-10
- [38] IBM Institute of Business Value IBM, (2018), “Analytics: The real-world use of big data in retail”, How innovative retailers extract value from uncertain data, <https://www-935.ibm.com/services/us/gbs/thoughtleadership/big-data-retail/>

- [39] IDC, (2014), "Executive Summary: Data Growth, Business Opportunities, and the IT Imperatives", Retrieved from <https://www.emc.com/leadership/digital-universe/2014iview/executivesummary.html>
- [40] Integreon Insight (2012). "Big just got bigger". Grail Research, 1-17 (Retrieved from http://www.integreon.com/pdf/Blog/Grail-Research-Big-Data-Just-Got-Bigger_232.pdf)
- [41] Jager, J., Putnick, D. L., & Bornstein, M. H. (2017). II. More than just convenient: The scientific merits of homogeneous convenience samples. *Monographs of the Society for Research in Child Development*, 82(2), 13-30
- [42] Jun, S. P., & Park, D. H. (2017). Visualization of brand positioning based on consumer web search information: Using social network analysis. *Internet Research*, 27(2), 381-407.
- [43] Kaur, R., & Jagdev, G. (2017). Big Data in retail sector-an evolution that turned to a revolution. *International Journal of Research Studies in Computer Science and Engineering (IJRSCSE)*, 4(4), 43-52
- [44] Khan, A., (2017, July 31), "Trading data", Retrieved from <https://www.thenews.com.pk/magazine/money-matters/194614-Trading-data>
- [45] Khan, M. W., Khan, M. A., Alam, M., & Ali, W. (2018). Impact of Big Data over Telecom Industry. *Pakistan Journal of Engineering, Technology & Science*, 6(2), 116-126, <http://dx.doi.org/10.22555/pjets.v6i2.1958>
- [46] Kivunja, C., & Kuyini, A. B. (2017). Understanding and applying research paradigms in educational contexts. *International Journal of Higher Education*, 6(5), 26-41
- [47] Kshetri, N. (2016). *Big data's big potential in developing economies: impact on agriculture, health and environmental security*. CABI. Wallingford, doi10.1079/9781780648682.0000, ISBN 9781780648682
- [48] Latif, Z., Tunio, M. Z., Pathan, Z. H., Jianqiu, Z., Ximei, L., & Sadozai, S. K. (2018, March). A review of policies concerning development of big data industry in Pakistan: Subtitle: Development of big data industry in Pakistan. In *Computing, Mathematics and Engineering Technologies (iCoMET)*, 2018 International Conference on (pp. 1-5). IEEE.
- [49] Le, T. M., & Liaw, S. Y. (2017). Effects of Pros and Cons of Applying Big Data Analytics to Consumers' Responses in an E-Commerce Context. *Sustainability*, 9(5), 798.
- [50] Leedy, P. D., & Ormrod, J. E. (2005). *Practical Research Planning and Design*. New Jersey: Pearson Merrill Prentice Hall.
- [51] Lycett, M. (2013). 'Datafication': making sense of (big) data in a complex world. *European Journal of Information Systems*, 22(4), 381-386
- [52] Maheshwari, A. (2014). *Business Intelligence and Data Mining*. Business Expert Press.
- [53] Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey & Company, Washington, USA,

- [54] Matsa, D. A. (2011) Competition and product quality in the supermarket industry, *The Quarterly Journal of Economics*, 126, 1539–91. doi:10.1093/qje/qjr031
- [55] McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: the management revolution. *Harvard business review*, 90(10), 60-68.
- [56] Mkansi, M., & Acheampong, E. A. (2012). Research philosophy debates and classifications: students' dilemma. *Electronic journal of business research methods*, 10(2), 132-140
- [57] Mugenda, A. (2003). *Research methods Quantitative and qualitative approaches* by Mugenda. Nairobi, Kenya.
- [58] Mukherjee, S., & Shaw, R. (2016). Big Data–Concepts, Applications, Challenges and Future Scope. *International Journal of Advanced Research in Computer and Communication Engineering*, 5(2), 66-74.
- [59] New Desk. (2020, May 13). Profit, <https://profit.pakistantoday.com.pk/2020/05/13/pakistan-retains-its-place-in-msci-emerging-markets-index/>
- [60] Oracle (2012). Big data for the enterprise. Oracle White Paper, 1–14, <http://www.oracle.com/us/products/database/big-data-forenterprise-519135.pdf>.
- [61] Oso, W. Y., & Onen, D. (2009). *A general guide to writing research proposal and report*. Jomo Kenyatta Foundation.
- [62] Pantano, E., Giglio, S., & Dennis, C. (2019). Making sense of consumers' tweets. *International Journal of Retail & Distribution Management*, 47(9), 915-927.
- [63] Pathirage, C. P., Amaratunga, R. D. G., & Haigh, R. P. (2008). The role of philosophical context in the development of research methodology and theory. *The Built and Human Environment Review*, 1(1), 1-10
- [64] Pecori, R. (2016). S-Kademlia: A trust and reputation method to mitigate a Sybil attack in Kademlia. *Computer Networks*, 94, 205-218.
- [65] Ravand, H., & Baghaei, P. (2016). Partial least squares structural equation modeling with R. *Practical Assessment, Research, and Evaluation*, 21(1), 1-16, <https://doi.org/10.7275/d2fa-qv48>
- [66] Rejeb, A., Rejeb, K., & Keogh, J. G. (2020). Potential of Big Data for Marketing: A Literature Review. *Management Research and Practice*, 12(3), 60-73
- [67] Ridge, M., Johnston, K. A., & O'Donovan, B. (2015). The use of big data analytics in the retail industries in South Africa. *African Journal of Business Management*, 9(19), 688-703
- [68] Salvador, A. B., & Ikeda, A. A. (2014). Big data usage in the marketing information system. *Journal of Data Analysis and Information Processing*
- [69] Santoro, G., Fiano, F., Bertoldi, B., & Ciampi, F. (2019). Big data for business management in the retail industry. *Management Decision*, 57(8), 980-1992
- [70] Saunders, M. N. K., Lewis, P., Thornhill, A., & Bristow, A. (2015). Understanding research philosophies and approaches: Research methods for business students

- [71] Saunders, M., Lewis, P. & Thornhill, A. (2007). Research methods. Business Students 4th edition Pearson Education Limited, England
- [72] Schultz, J., (2017, October 10), "How Much Data is Created on the Internet Each Day?" <https://blog.microfocus.com/how-much-data-is-created-onthe-internet-each-day>
- [73] Seetharaman, A., Niranjan, I., Tandon, V. and Saravanan, , A. S., (2016), "Impact of Big Data on the Retail Industry" Journal of Corporate Ownership & Control, 14(1), 506-518
- [74] Sekaran, U. and Bougie, R., (2016), "Research Methods For Business: A Skill Building Approach", John Wiley & Sons, 1-448, ISBN 1119165555, 9781119165552
- [75] Shankar, V. (2019). Big Data and Analytics in Retailing. NIM Marketing Intelligence Review, 11(1), 36-40
- [76] Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. Psychometrika, 74(1), 107
- [77] Sijtsma, K. (2009). Over misverstanden rond Cronbachs alfa en de wenselijkheid van alternatieven. Psycholoog, 44(11), 561
- [78] Sileyew, K. J. (2019). Research Design and Methodology. In Text Mining-Analysis, Programming and Application. Intech Open, <https://www.intechopen.com/books/cyberspace/research-design-and-methodology>
- [79] Spiess, J., T'Joens, Y., Dragnea, R., Spencer, P., & Philippart, L. (2014). Using big data to improve customer experience and business performance. Bell labs technical journal, 18(4), 3-17.
- [80] State Bank of Pakistan, (2014), "'Supermarkets and Retail Shops", Research Report on 'Supermarkets and Retail Shops' Segment", http://www.sbp.org.pk/departments/ihfd/Sub_Segment%20Booklets/Supermarkets%20and%20Retail%20Shops.pdf
- [81] Stoicescu, C. (2016). Big Data, the perfect instrument to study today's consumer behavior. Database Syst. J, 6, 28-42.
- [82] Surbakti, F. P. S., Wang, W., Indulska, M., & Sadiq, S. (2020). Factors influencing effective use of big data: A research framework. Information & Management, 57(1), 103146
- [83] Tankard, C. (2012). Big Data Security. Network Security Newsletter, Elsevier, ISSN 1353-4858
- [84] Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. International journal of medical education, 2, 53-55, doi 10.5116/ijme.4dfb.8dfd
- [85] Thakuriah, P. V., Tilahun, N. Y., & Zellner, M. (2017). Big data and urban informatics: innovations and challenges to urban planning and knowledge discovery. In Seeing cities through big data (pp. 11-45). Springer, Cham.
- [86] Thau, B., (2016, March 2), "Retail pricing strategies getting a makeover from data analytics", IBM Big Data and Analytics Hub, <https://www.ibmbigdatahub.com/blog/retail-pricing-strategies-getting-makeover-data-analytics>
- [87] Valchanov, I. (2017). Is data science really a rising career? <https://www.quora.com/Is-data-science-really-a-rising-career/answer/IliyaValchanov>

- [88] Vargas-Sánchez, A., do Valle, P. O., da Costa Mendes, J., & Silva, J. A. (2015). Residents' attitude and level of destination development: An international comparison. *Tourism Management*, 48, 199-210.
- [89] Vijayarani, S., & Sharmila, S. (2016, August). Comparative analysis of association rule mining algorithms. In *2016 International Conference on Inventive Computation Technologies (ICICT)* (Vol. 3, pp. 1-6). IEEE
- [90] Vikas, D. & Nadir, Z. (2014). *Big data and social media analytics*. A Cambridge Assessment publication, 18, 36-41, www.cambridgeassessment.org.uk/research-matters/
- [91] Voleti, S., Kopalle, P.K., & Ghosh, P. (2015). An inter-product competition model incorporating branding hierarchy and product similarities using store-level data. *Management Science*, 61(11), 2720-2738.
- [92] Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. *Journal of Marketing*, 80(6), 97-121.
- [93] Werner, K., Lerzan, A., Yakov, B., Kristina, H., Sertan, K., Francisco, V. O., Marianna S., David, D. and Babis, T., (2017) "Customer engagement in a Big Data world", *Journal of Services Marketing*, 31(2), 161-17
- [94] Zanini, N., & Dhawan, V. (2015). Text Mining: An introduction to theory and some applications. *Research Matters*, 19, 38-45
- [95] Zeng, J., & Glaister, K. W. (2018). Value creation from big data: Looking inside the black box. *Strategic Organization*, 16(2), 105-140
- [96] Zikopoulos, P., & Eaton, C. (2011). *Understanding big data: Analytics for enterprise class hadoop and streaming data*. McGraw-Hill Osborne Media. Erevelles, S., Nobuyuki
- [97] Žukauskas, P., Vveinhardt, J., & Andriukaitienė, R. (2018). Philosophy and paradigm of scientific research. *Management Culture and Corporate Social Responsibility*, 121.