Analysis on User Activity in E-Commerce Website for Performance Evaluation and Decision Making Using Big Data Analytics

Radityo Pradana[1], Rizky Luxianto[2]
{radityo_pradana@yahoo.com[1], rizky.lux@gmail.com[2]}

Universitas Indonesia[1,2]

Abstract. The purpose of this study is to analyze how Big Data Analytics can help managers in their decision making process, specifically in e-commerce / web-based business. This study took the dataset from one website that implements e-commerce functions in its website. The data taken would be the data of user activities in the website such as page visit, add product to cart, and buy online. To analyze the dataset, several algorithm will be used, such as Association Rule Mining Algorithm (APRIORI), K-Means Clustering, and Pearson's Correlation Coefficient. The data will be processed and will be used to show details in users behavior in the website to see the specific pattern that can be useful for decision making. For example, which product is visited and bought the most, how many pages are visited before buying the product, how many users repurchase the product, and which retailer is mostly used by users. This paper would also identify whether the current Big Data implementation on the company can be improved, and identify if it is a good investment for the company to improve the Big Data implementation system.

Keywords: Big Data; e-commerce; Website; Web-Based Business; Decision Making.

1. Introduction

Firms with the most capability to gain data and process them into useful and structured information which then be converted into sharp actions will most likely standout in the competition. Not only data from macro aspects and competitor's behavior, but also people's behavior to our products. Learning how people think, talk and react towards our product can be beneficial information that can tell us which action should be taken.

In order to do this, the firm needs to gain accurate information, not manipulated, spontaneous and candid from people. When people are manually asked or handed forms for them to fill, there are possibilities that what they write or answer does not represent their natural answer, as they might answer in accordance to their preference or benefit. This would cause irrelevant data collected and on extreme cases, causing firms to take wrong actions. Hence, what we are looking for is an "unconscious" answer, which can be achieved through Big Data.

As for its implementation in ecommerce / online retailing, the research towards online retailing has been growing substantially in many aspects. For example, online shoppers' purchase behavior by accounting for the sequence of pages viewed or tasks completed at a website (Montgomery, Alan, Shibo, Srinivasan, & Liechty, 2004), the cumulative effect of visits made between purchases [1], shoppers' cognitive style and website design [2], the effect of consumer reviews [3], the association of visit patterns

ICONISTECH-1 2019, July 11-12, Bandung, Indonesia

DOI 10.4108/eai.11-7-2019.2298019

and purchase behavior, and the effect of search refinement tools [4]. The growth of research in this aspect makes this research as one of the most trending discussions in this recent period.

2. Literature Review

Big data can be characterized by 5Vs Volume, Velocity, Variety, Veracity, and Variability, which will be explained below [5]:

1. Volume generally reflects the space required to store data.
2. Velocity reflects the speed of data transmission and processing, i.e., how effectively, and efficiently real-time data is collected and processed.
3. Variety reflects the type of data, i.e., data can be structured or unstructured and can also be in different forms such as text, image, audio, and video.
4. Veracity reflects the degree to which data can be trusted.
5. Variability reflects the dissimilarity between different instances in a data set.

Building a website essentially relies on one of the most important aspects of the user, which is its behavior. Talking about behavior, one of the most famous theories about what humans require in order to reach their "pinnacle" — the point where they decide to participate — was posited by the humanist psychologist Abraham Maslow in the form of a "hierarchy of needs."

William Craig, CEO of WebFx to explain what human prioritize and what they need in order to achieve “self-actualization”, sorted by its importance.

1. **Accessibility**. The website can be found and used by all people.
2. **Stability**. The website is consistent and trustworthy.
3. **Usability**. The website is user-friendly.
4. **Reliability**. The website is consistently available, without downtime.
5. **Functionality**. The website offers content, tools and services users’ value.
6. **Flexibility**. The website adapts to the needs and wants of users.

According to the model, it is important for a website to serve and prioritize what users think is important, especially when the website is an e-commerce website which goals is to generate sales. [6]

3. Research Methodology

The data that will be collected in this research is Big Data of web visitors’ activities that is recorded using a custom-made third-party data analytics tool. The Analytics tool catches and records activities of users within the website, and then present the collected data in forms of well-designed reports. The analytics tools for running the data will be created using node.js in a form of local website. The input will be the website activity in 2018.

After determining the components of the analysis, the analysis will be executed by steps as pictured in the graph below.

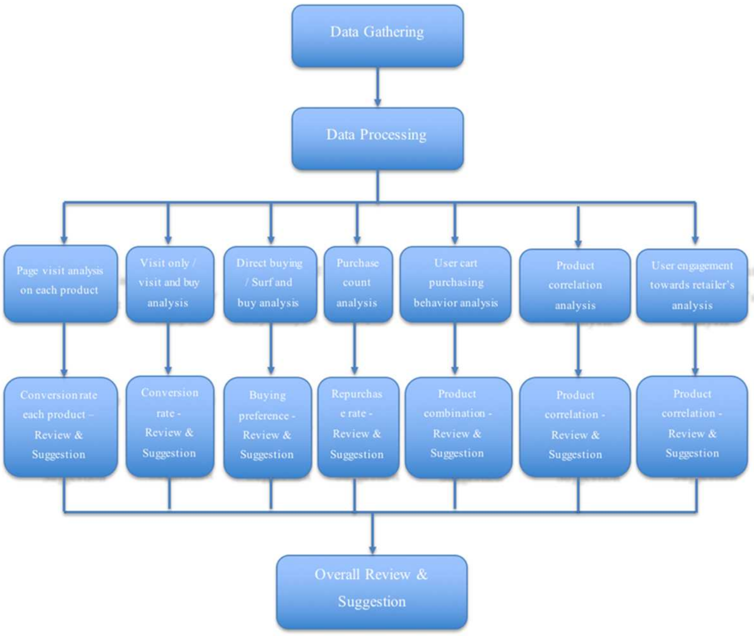

Figure 1 Analysis

The first analysis of this research is to see the users page visit behavior in the website. For each product URL, the analysis will first search activities that is tagged as buy_online and page_visit. On each activity, the session_id (user ID) will be counted, and the number of session_id that engages page visit (without buy_online) and the number of session_id that engages page visit and also buy_online will be reported. The report will be summarized every quarter year and the percentage of buy_online and visit will be provided. We will see this number as "conversion rate for each product URL", which is calculated as follow.

Conversion Rate = buy_online / page_visit

The second analysis of this research will count the overall conversion rate for the website. Instead of analyzing the conversion rate for each product (from Product Detail Page / PDP), this analysis will also count the other components of the website, such as home page, product category page, etc.

The third analysis of this research will analyze the buying behavior of the user. Specifically, this analysis will see the user's movement before buying, whether the user prefers to surf the website before buying (opening more than 4 pages before buying) which will be marked as "surf and buy" consumer, or directly buy the product (opening 4 or less pages before buying) which will be marked as "direct buyer". The application will count the number of pages opened by a unique user before proceeding to purchase the product.

The fourth analysis of this research will analyze the buying count for each unique user. Essentially, we will count how many times does each unique user use the buy online feature. The data field that will be used for this research is session_id and buy_online activities.

In the result of the analysis, "Repurchase Rate" can then be seen. As one of the most used KPI in ecommerce, Repurchase Rate represents the rate of consumer's loyalty to the product. Repurchase Rate is calculated using the following formula [7]

RR = ((PC1*O1+PC2*O2+PC3*O3+...+PCn*On) / Total Transaction)

PC = Purchase Count

O = Occurrence

From Repurchase rate generated from the analysis, we will then compare the repurchase rate of XYZ website and other ecommerce website, to evaluate the performance of the business. From the evaluation, suggestion can then be proposed based on the analysis.

The fifth analysis of this research will seek the top combination of two products sold by the business. Essentially, we will analyze each cart that is forwarded to retailers' website, and see which product combinations that are frequently bought together in one cart. Using Association Rule Mining / Apriori Algorithm, which is one common method used in Machine Learning. The steps towards the datasets are as follows:

Step 1: Collect all carts sent with more than one product
Step 2: Filter product bought count more than Support Threshold
Step 3: From Apriori Array 1, generate combination of 2 products.
Step 4: Find occurrence in all combinations defined in step 3 that passes the support threshold

The sixth analysis of this research is to find correlation between products, and find which products have strong correlation between each other. For this, **Pearson's Correlation** will be used.

Step 1: Collect all sales data
Step 2: Find all combinations of the 5 top products
Step 3: Find correlation value of each product combination

The seventh analysis of this research is to find each retailers performance in the website. As XYZ website utilizes online retailers such as Walmart, Walgreens, and Amazon to provide users with retailer channels, we will look into the performance of each retailer, so that user's preferences / behavior towards retailer can be analyzed.

For this analysis, we will record all buy_online activities, the sales in the activity, and which retailer is used for each buying activity. After the data has been processed, sales report for each retailer will then be generated and evaluated, and suggestion according to the report will be proposed.

4. Results

The first analysis towards the dataset is the page visit of the website. This analysis is done to analyze the user engagement level towards each product, specifically to each PDP.

URL	Q1			Q2			Q3			Q4			total		
	visit	buy	percentage	visit	buy	percentage	visit	buy	percentage	visit	buy	percentage	visit	buy	percentage
Product 28	1005	21	2.09%	1235	57	4.62%	1195	45	3.77%	607	20	3.29%	4042	143	3.54%
Product 37	4263	91	2.13%	4950	102	2.06%	5909	173	2.93%	2415	101	4.18%	17537	467	2.66%
Product 36	19564	367	1.88%	18202	384	2.11%	25269	780	3.09%	12474	451	3.62%	75509	1982	2.62%
Product 11	3306	85	2.57%	4377	94	2.15%	4127	110	2.67%	1467	48	3.27%	13277	337	2.54%
Product 22	5695	108	1.90%	6520	143	2.19%	6302	161	2.55%	2710	125	4.61%	21227	537	2.53%
Product 49	4021	88	2.19%	4448	98	2.20%	4841	115	2.38%	2294	81	3.53%	15604	382	2.45%
Product 48	8510	130	1.53%	9002	146	1.62%	10965	238	2.17%	4391	176	4.01%	32868	690	2.10%
Product 46	3120	32	1.03%	3450	62	1.80%	5608	128	2.28%	2696	77	2.86%	14874	299	2.01%
Product 21	8908	140	1.57%	10040	186	1.85%	12343	245	1.98%	5713	148	2.59%	37004	719	1.94%
Product 3	364	6	1.65%	589	12	2.04%	703	13	1.85%	258	6	2.33%	1914	37	1.93%
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Product 5	6203	5	0.08%	7271	13	0.18%	8613	14	0.16%	2172	6	0.28%	24259	38	0.16%
Total	402062	3725	48.06%	479445	4301	52.15%	494704	5743	61.56%	201758	3308	90.58%	1577969	17077	1.08%
Max Value	87901	818	2.57%	95982	769	4.62%	98573	980	3.77%	51343	471	4.61%	333799	3038	3.54%
Min Value	364	3	0.08%	589	8	0.17%	703	9	0.16%	258	4	0.18%	1914	26	0.16%

Figure. 2. First Analsysis

The analysis on page visits show which product URL that is mostly visited in the period. This of course represents the performance of the product itself. Other than that, the analysis also shows the ratio between the visit rate and buy rate, known as the "conversion rate" for each product.

In the result, it can be seen the maximum conversion rate is 3.54% (product 28), and the minimum conversion rate is 0.16% (product 5). As the number difference is rather high (1 buyer for every 28 visitors vs 1 buyer for every 638 visitors), one of the suggestions that can be made is to increase the conversion rate for product 5.

The second analysis is consumer behavior to see the percentage of users who only visit and who also buy the product, recapped quarterly. Below is the result:

	Q1	Q2	Q3	Q4	Total	Max	Min
Visit Only	589,779.00	899,924.00	829,369.00	348,600.00	2,667,672.00	899,924.00	348,600.00
Visit and Buy	3,203.00	4,930.00	5,986.00	3,280.00	17,399.00	5,986.00	3,203.00
Percentage	0.54%	0.55%	0.72%	0.94%	0.65%	0.94%	0.54%

Figure 3. Second Analysis

In further analysis of this data, we would like to measure the effectiveness of the website in terms of its conversion rate. Below is the data of average conversion rate from 2015 to 2018.

Average Website Conversion Rate

Quarter	Global	US	UK
Q4 2014	3.42%	3.60%	3.65%
Q1 2015	2.83%	2.91%	3.56%
Q2 2015	3.08%	3.18%	3.97%
Q3 2015	3.02%	3.09%	4.08%
Q4 2015	3.48%	3.62%	4.20%
Q1 2016	2.94%	2.89%	5.01%
Q2 2016	2.76%	2.72%	4.64%
Q3 2016	2.46%	2.44%	4.35%
Q4 2016	2.95%	3.00%	4.45%
Q1 2017	2.48%	2.46%	3.57%
Q2 2017	2.86%	2.62%	4.28%
Q3 2017	2.80%	2.56%	4.21%
Q4 2017	3.15%	2.96%	4.27%
Q1 2018	2.77%	2.60%	3.91%
Q2 2018	2.86%	2.63%	4.31%

Figure 4. Average Website Conversion Rate

From Q4 2014 to 2018, it can be seen that in the US, the minimum conversion rate is 2.44% in Q3 2016, and the maximum conversion rate is 3.62% Q4 2015. In 2018, the conversion rate is around 2.60%. This data will be used as benchmarks for the data of XYZ Corporation.

The third analysis would analyze the consumer behavior in the website by counting how many page visits are executed by user before proceeding to buy the product. User who surfs (visits more than 4 pages before buying the product) tends to see the information first before buying the product, while direct buyers (4 or less visits before buying) tends to have enough information about the product, hence directly accessing the PDP and purchase the product.

The analysis starts by using k-means analytics method to put the data into clusters, below is the result of the clustering.

	Low	High	min	Cluster
Q1	19,834.22	827,845.00	19,834.22	low
Q2	41,030.89	445,633.00	41,030.89	low
Q3	689,259.89	0.00	0.00	high
Q4	42,682.22	897,849.00	42,682.22	low

Figure 5. Using K-Means

Following the clustering, the data is then divided into two categories, which is direct buy and surf and buy (non-direct buy). It is divided into such categories to link users' activities with the website's conversion path. The result of this analysis can then determine whether the conversion path designed by the website is effective or not.

	Q1	Q2	Q3	Q4	Total	Max	Min
Direct buy	2,542.00	3,141.00	4,372.00	2,525.00	12,580.00	4,372.00	2,525.00
Surf and buy	1,852.00	1,790.00	2,254.00	1,289.00	7,185.00	2,254.00	1,289.00
Percentage	72.86%	56.99%	51.56%	51.05%	57.11%	72.86%	51.05%

Figure 6. Result Analysis

It shows that 57.11% of user visit less than 4 pages prior to buying the product. From one point of view, it can be viewed that people tend to not spend more time to surf around and prefer to buy the product directly. However, looking at the previous analysis where it shows that the conversion rate is rather low, one assumption to be made is that when the steps to buying the product is long, people tend to cease buying the product. The evidence is rather relevant by looking at this data. Further analysis on this assumption will be executed by looking at the consumers "road" towards buying the product.

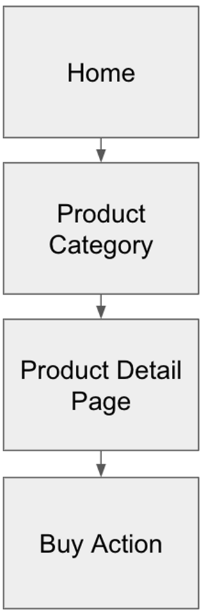

Figure 7. Consumers

The graph shows the 4 consumers step that user needs to take in order to purchase the product in the website. The data on third analysis implies that in terms of buying the product, users prefers less step in buying the product. This can mean that users might already have enough information about the product and prefers to buy the product within less steps.

Users prefer direct buying. The 4 steps to product purchase in the current XYZ website can be improved to follow this user behavior. For example, by shortening the step to product purchase. This can be done by having the add to cart button in product category page, or having direct linkage to product detail page from the home page. As following the user behavior, it would also potentially increase the conversion rate.

The fourth analysis analyzes how many times unique users proceed to purchase the product. below is the result:

Purchase Count	Occurrence
1	11557
2	2919
3	187
4	205
5	18
6	29
7	3
8	5
9	1
10	1
11	1
14	1
18	1
66	1

Number of Buy	19229
Number of Unique User	14929

Figure 8. Result

Looking at the results, it can be seen that most of purchase action is executed only one time. This means that users who chooses to do repurchase is rather small. However, with the data, we can then calculate one of the benchmarks that can be used to measure the purchase behavior, that is, the Repurchase Rate.

RR = ((PC1*O1+PC2*O2+PC3*O3+...+PCn*On) / Total Transaction)

Using the formula, the calculated repurchase rate of XYZ website in 2018 is 39.9%.

To evaluate this number, we will need to look at the repurchase rate benchmark in the market. Alex Schultz, VP of Growth at Facebook, in its content “How to Start a Startup”, mentions that “If you can get 20–30% of your customers coming back every month and making a purchase from you, then you should do pretty well.”. when this statement is used as benchmark for data result of XYZ website, then it shows that the repurchase rate of XYZ company is rather high for the market.

From one point of view, it can be assumed that the consumer who buys the product once is likely to be satisfied with the product and continue to use the product again. This is a good sign since even though the previous analysis shows that the conversion rate is low, once the consumer buys the product, they proceed to repurchase the product.

Repurchase rate is high compared to average ecommerce conversion rate in the US. As we understand from this analysis that customer likely subscribe to the product, XYZ company can act to improve this condition, for example by having more personalized offers towards repurchasing customer, such as promotion on the 2nd buy, customer service to contain customers feedback, etc.

The fifth analysis looks more into user behavior, specifically in purchasing action. The objective of this analysis is to find percentage of product which are frequently bought together in one cart.

The result of the analysis will provide top 2-product-combination which are frequently bought together in the cart.

Item One	Item Two	Count
product 4	product 8	45
product 7	product 8	32
product 20	product 24	28
product 23	product 25	32
product 29	product 45	34

Figure 9. Result Analysis

This result can then be provided to the management to consider action that might increase user engagement. For example, making promotion of bundle purchase for the product combination, adding the "frequently bought together" feature, and advertising Prod B in Prod A page.

The objective of the sixth analysis is to find correlation between products in the website. In this case, we will first collect all sales data and create monthly sales report for each product available. Then we will choose 5 top products with most sales. After that, all combinations within the top products will be created and will be used for next calculation for the correlation. To find correlations between product, **Pearson's Correlation** value will be used.

From the generated sales report, 5 top products are selected and all combination will be defined. Then, Pearson's Correlation value between each combination will be calculated. The result is as shown below:

Prod A	Prod B	Correlation
product 3	product 8	0.734
product 3	product 23	0.7878
product 3	product 20	0.7235
product 3	product 4	0.7358
product 8	product 23	0.8337
product 8	product 20	0.8717
product 8	product 4	0.8728
product 23	product 20	0.877
product 23	product 4	0.782
product 20	product 4	0.8851

Fi gure 10. Result Analysis

From the result, management can see which sales movement of two products that have strong correlation. This information can then be used to make decision making. Such as having more promotion for the sales of 2 certain products.

The seventh analysis comes as the website does not possess login and payment gateway feature, it instead utilizes the functionality of retailer's API. The retailer's API allows website to redirect user to the website retailer with product already added in the retailer's cart, for user convenience to proceed to purchasing the product. Hence, we would also like to see the user engagement towards the retailer, to see which retailer gives most profit, which retailer gives less, etc. Below are the results:

	Walmart	Amazon	Walgreens	Total Quarterly Sales
Q1	56,531.80	44,485.14	0.00	101,016.94
Q2	53,929.32	50,537.85	0.00	104,467.17
Q3	80,103.22	59,476.04	443.38	140,022.64
Q4	45,746.46	26,546.89	3,727.15	76,020.50
Total sales each retailer	236,310.80	181,045.92	4,170.53	421,527.25

Figure 11. Result

With sales report by retailers created, management can determine which retailer gives the most profit and which one gives less. This can be interpreted into decision such as increasing engagement for the retailers that did not perform very well, or even discontinue the usage of the retailer. For most used retailer, a push for advertisement can be done to increase user acknowledgement towards the product and the service.

5. Discussion

The summary of the research will be described in points according to each analysis. From each analysis, the summary of the result is as follows:

1. From page visit analysis on each product, it provides the list of product performance, indicated by their conversion rate. This can help management to analyze the performance of each product, by seeing which product has high conversion rate, and which is low. The suggestion to this analysis is to increase the conversion rate of the current low-conversion rate product, by implementing the literature.
2. From website page visit analysis, it provides the overall performance of the website, indicated by its overall conversion rate. As one of the most used KPI in ecommerce, conversion rate of XYZ company represents the effectiveness of the website. From the analysis result, the conversion rate of XYZ website is rather low, and steps according to literature need to be executed to increase its conversion rate.
3. From direct and surf and buy analysis, the result complements the conversion rate value that is found in the previous analysis. From the result, it can be seen that user prefers to directly buying the product rather than surfing before buying. Knowing this, XYZ can see one way to potentially increase the conversion rate by shortening the path to conversion.
4. From purchase count analysis, it creates another one of the most used KPIs in ecommerce, which is repurchase rate. The repurchase rate represents customers loyalty of the product. In the analysis result, the repurchase rate of the website is quite high, which means the customer of the product is rather satisfied and tend to continue buying the product. This is a good sign for XYZ company. In summary, XYZ company should seek for opportunities to obtain new customers, as the customers would be likely to subscribe to the product, once they know and try to use the product.
5. The user cart purchasing behavior analysis shows managers which product is likely to be bought together in one cart. The result of this analysis is rather useful, as it can help managers to acknowledge such product combinations, and use the opportunity by implementing marketing strategy, such as bundling product, promotion, etc.
6. The product correlation analysis shows which product has strong correlation in terms of sales. From one point, it can be seen that the products with similar correlation have identical sales movement in one-year period. From this, managers can then see which product has high sales in which quarter, and decide the push marketing in correct period of the year.
7. The retailer engagement analysis shows which retailer provides most profit for the website. For manager, this analysis is useful as it can see the performance of each retailer, and create movement / strategy according to the result. According to analysis result, Walmart retailer provides most sales comparing to the other retailers. This can be one point of consideration for managers for which action to be taken. For example, to analyze which factor determines the low sales in the low-selling retailer, and push / increase the sales for the retailer.

The Big Data implementation that is done by XYZ website is proven to be able to provide useful information for managers, looking at values and analysis that can be created from the dataset.

6. Conclusion

The overall research concludes that the analysis of big data analytics can significantly help managers in seeing the performance of the company website. The result that can be extracted from the dataset can be very helpful for managers to see how the ecommerce is going, and which action that needs to be taken according to the analysis result. This research gives merit to the industry that it gives example and insight on how they can improve their decision making with the help of Big Data Analytics, and how it can be used to gain competitive advantage in the process.

References

[1] W. W. Moe dan S. Peter, "Dynamic Conversion Behavior at E-Commerce Sites," *Management Science,* p. 326–35, 2004.

[2] J. R. Hauser, U. G. L, G. Liberali dan M. Braun, "Website Morphing," *Marketing Science,* p. 202–23, 2009.

[3] F. Zhu dan X. Zhang, "Impact of online consumer reviews on sales: the moderating role of product and consumer characteristics," *J. Mark,* p. 133–148, 2010.

[4] Y. Chen dan S. Yao, "Sequential Search with Refinement: Model and Application with Click-Stream Data," *Management Science,* 2016.

[5] H. Liu dan M. Cocea, Granular Computing Based Machine Learning: A Big Data Processing Approach, Cham: Springer, 2018.

[6] W. Craig, "Human Behavior Theories That Can be Applied to Web Design," 2019. [Online]. Available: https://www.webfx.com.

[7] S. Matteo, "Repurchase rate—the most overlooked eCommerce KPI," June 2016. [Online]. Available: https://medium.com.

[8] F. Zhu dan M. Zhang, "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics," *Journal of Marketing,* p. 133–48, 2010.

[9] A. Yasmin, S. Tasneem dan K. Fatema, "Effectiveness of Digital Marketing in the Challenging Age: An Empirical Study," *International Journal of Management Science and Business Administration, Volume 1, Issue 5,* pp. 69-80, 2015.

[10] G. Udo dan G. Marquis, "Factors affecting e-commerce web site effectiveness," *J. Comput. Inf. Syst,* p. 10–16, 2002.

[11] A. Stathopoulou dan G. Balabanis, "The effects of loyalty programs on customer satisfaction, trust, and loyalty toward high-and low-end fashion retailers," *J. Bus. Res,* p. 5801–5808, 2016.

[12] M. R. Solomon, Consumer Behavior. Buying, Having, and Being, Harlow: Pearson, 2018.

[13] C. Simoni, "L'impatto del commercio elettronico business-to-consumer sulla logistica distributiva, con particolare riferimento alla logistica dell'"ultimo miglio"," *Sinergie Ital. J. Manag,* p. 151–175, 2011.

[14] P. Selvidge, "How long is too long to wait for a website to load," *Usability News 1,* p. 1–3, 1999.

[15] P. Sandeen, "Make Buying Inevitable," 2019. [Online]. Available: https://neilpatel.com/.

[16] M. Salehi, H. Mirzaei, M. Aghaei dan M. Abyari, "Dissimilarity of e-marketing vs traditional marketing," *Int ernational Journal of Academic Research in Business and Social Science,* pp. 510-515, 2012.

[17] R. Pozharliev, W. Verbeke, J. V. Strien dan R. Bagozzi, "Merely being with you increases my attention to luxury products: using EEG to understand consumers' emotional experience with luxury branded products," *J. Mark. Res,* p. 546–558, 2015.

[18] B. Plaza, "Google analytics for measuring website performance," *Tour. Manag,* p. 477–481, 2011.

[19] J. Petersen dan V. Kumar, "Perceived risk, product returns, and optimal resource allocation: evidence from a field experiment," *J. Mark,* p. 268–285, 2015.

[20] S. Peter dan Y.-H. Park, "Browsing Behavior at Multiple Websites," *Marketing Science,* p. 280–303, 2004.

[21] S. Peinkofer, T. Esper, R. Smith dan B. Williams, "Assessing the impact of price promotions on consumer response to online stockouts," *J. Bus. Logist,* p. 260–272, 2015.

[22] M. Pasaribu, Big Data and Strategic Entrepreneurship, Jakarta: NT Corp, 2017.

[23] R. Owen dan P. Humphrey, "The structure of online marketing communication channels," *Journal of Management and Marketing Research,* pp. 1-10, 2008.

[24] N. Ostapenko, "Online discount luxury: In search of guilty customers," *Int. J. Bus. Social. Res,* p. 60–68, 2013.

[25] A. Nerurkar, "Online marketing - challenges in future," *International Journal of Research in Humanities, Arts and Literature,* pp. 183-188, 2014.

[26] Montgomery, L. Alan, L. Shibo, K. Srinivasan dan J. C. Liechty, "Modeling Online Browsing and Path Analysis Using Clickstream Data," *Marketing Science,* p. 579–95, 2004.

[27] K. Matthews, "5 Industries Becoming Defined by Big Data and Analytics," August 2018. [Online]. Available: https://towardsdatascience.com.

[28] R. Ladhari, "Developing e-service quality scales: a literature review," *J. Retail. Consum. Serv,* p. 464–477, 2010.

[29] Z. L, S. A, B. M, M. S, S. T, P. R, W. S, L. H dan K. D, "Visual analytics for the big data era—a comparative review of state-of-the-art commercial systems," *Proceedings of the IEEE Conference on Visual Analytics Science and Technology,* p. 173–182, 2012.

[30] Koukova, N.T., J. Srivastava dan S.-F. M, "The effect of shipping fee structure on consumers' online evaluations and choice," *J. Acad. Mark. Sci,* p. 759–770, 2012.

[31] M. Kiang, T. Raghu dan Shang, "Marketing on the Internet — who can benefit from an online marketing approach?," *Decision Support Systems,* p. 383–393, 2000.

[32] T. Jackson, The process of trend development leading to a fashion season. Fashion Marketing: Contemporary Issues, Oxford: Elsevier, 2007.

[33] B. Hernández, J. Jiménez dan M. Martín, "Customer behavior in electronic commerce: the moderating effect of e-purchasing experience," *J. Bus. Res,* p. 964–971, 2010.

[34] R. Gautam, "Internet Marketing Usage by Small Indian Entrepreneurs: An Exploratory Study of Punjab," *International Review of Management and Marketing, Vol. 2, No. 1,* pp. 43-51, 2000.

[35] M. J. Garbade, "Understanding K-means Clustering in Machine Learning," 12 September 2018. [Online]. Available: https://towardsdatascience.com.

[36] A. Gandomi dan M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," *International Journal of Information Management,* pp. 137-144, 2014.

[37] D. Fatta, R. Musotto dan V. R, "Analyzing e-commerce websites: a qualiquantitative approach for the User Perceived Web Quality (UPWQ)," *Int. J. Mark. Stud,* p. 33–44, 2016.

[38] D. Fagella, "What is Machine Learning?," June 2018. [Online]. Available: https://www.techemergence.com/.

[39] Fader, S. Peter, G. Bruce, Hardie dan H. Chun-Yao, "A Dynamic Changepoint Model for New Product Sales Forecasting," *Marketing Science,* p. 280–303, 2004.

[40] N. Essaides, "Five Key Benefits Big Data Can Deliver For Finance: Part 2," March 2016. [Online]. Available: https://www.digitalistmag.com.

[41] M. Epstein dan K. J. Yuthas, Evaluating the Effectiveness of Internet Marketing Initiatives, The Society of Management Accountants of Canada, 2007.

[42] H. El-Gohary, "E-Marketing - A literature Review from a Small Businesses perspective," *International Journal of Business and Social Science, Vol. 1 No. 1,* p. 214 – 244, 2010.

[43] A. Eisenberg, "7 Ways Fintechs Use Machine Learning to Outsmart the Competition," 2016. [Online]. Available: https://igniteoutsourcing.com/.

[44] S. Chand, "Top 10 Important Nature/Characteristics of Business," [Online]. Available: http://www.yourarticlelibrary.com/business/top-10-important-naturecharacteristics-of-business/7494.

[45] D. Chaffey dan P. Smith, “E-marketing Excellence: planning and optimizing your digital marketing,” *Routledge.*

[46] A. Bower dan M. III, “Return shipping policies of online retailers: normative assumptions and the long-term consequences of fee and free returns,” *J. Mark,* p. 110–124, 2012.

[47] Becerril-Arreola, M. Leng dan M. Parlar, “Online retailers' promotional pricing, free-shipping threshold, and inventory decisions: a simulation-based analysis,” *Eur. J. Oper. Res,* p. 272–283, 2013.

[48] I. Aldridge, “What Can You Do With Big Data in Finance?,” March 2017. [Online]. Available: https://www.huffingtonpost.com.

[49] C. A, S. IY dan D. KC, “Analytics over large-scale multidimensional data: The big data revolution!,” *Proceedings of the ACM International Workshop on Data Warehousing and OLAP,* p. 101–104, 2011.

[50] STA Group, “WHAT IS BUSINESS ARCHITECTURE?,” 2019. [Online]. Available: http://www.stagrp.com/architecture/business-architecture/what-is-business-architecture/.

[51] Akamai Technologies, “The State of Online Retail Performance,” Akamai, 2017.

[52] Connect2India, “PHOTON INTERACTIVE PRIVATE LIMITED,” 2018. [Online]. Available: https://connect2india.com/PHOTON-INTERACTIVE-PRIVATE-LIMITED/2015714.

[53] Photon Interactive, “Photon - Who we are,” 2019. [Online]. Available: https://www.photon.in/who-we-are.

[54] BusinessDictionary, “Innovation Definition,” 2019. [Online]. Available: http://www.businessdictionary.com/definition/innovation.html.

[55] Photon, “Company Overview of Photon Infotech Private Limited,” 24 May 2019. [Online]. Available: https://www.bloomberg.com/research/stocks/private/people.asp?privcapId=53977579.