IAIC Transactions on Sustainable Digital Innovation (ITSDI)
Vol. 4 No. 2 April 2023

p-ISSN: 2686-6285
e-ISSN: 2715-0461

Big Data and Strengthening MSMEs After the Covid-19 Pandemic (*Development Studies on Batik MSMEs in East Java*)

Rumanintya Lisaria Putri[1], Sarif Hidayat[2], Suprapto[3], Eko Wahyono[4], Sriyono[5] Lili Rahmawati[6]
Research Center for Cooperative, Corporation and People's Economy[1,2,3], Research Center for Social Village and Conectivity[4], Post Graduate Program Study[5], Faculty of Economics and Business[6]
National Research and Innovation Agency, Indonesia [1,2,3,4]
University of Muhammadiyah Sidoarjo[5]
Brawijaya University[6]
Gedung B.J Habibie, Jl. M.H Thamrin No. 8 Jakarta Pusat 10340[1,2,3,4]
Jl. Mojopahit No. 666 B, Sidowayah, Celep Kec. Sidoarjo, Kabupaten Sidoarjo, Jawa Timur 61215[5]
Jl. Veteran, Ketawanggede, Kec. Lowokwaru, Kota Malang, Jawa Timur 65145[6]

e-mail: rumanintya.lisaria.putri@brin.go.id[1], sari003@brin.go.id[2] , supr039@brin.go.id[3] , ekow005@brin.go.id[4] , sriyono@umsida.ac.id[5] , lilrahmawati@student.ub.ac.id[6]

Author Notification
10 December 2022
Final Revised
21 December 2022
Published
10 January 2023

Putri, R. L., Hidayat, S. ., Suprapto, Wahyono, E. ., Sriyono, & Rahmawati, L. (2023). Big Data and Strengthening MSMEs After the Covid-19 Pandemic (Development Studies on Batik MSMEs in East Java). *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, *4*(2), 83–100. https://doi.org/10.34306/itsdi.v4i2.574

Abstract

This research aims to analyze the effect of online business on economic growth in Java Island, segment and optimize aspects that support the development of internet business in the locus of Java Island, especially East Java Province, and complement it with a The existence of social. The method used is a combination of 5 techniques, namely regression analysis, thematic map visualization, clustering, spatial analysis, and text mining. This research has major implications for the development of MSMEs because MSMEs have succeeded in implementing e-commers in their sales and this greatly affects economic growth. Through the support of other analyses, it will also be known that the factors that support economic growth are more complete. The results show that online business affects the economic growth of provinces in Java. Online business businesses need to continue to be optimized by improving the quality of HDI and web networks, especially in the development priority areas of East Java Province. Programs related to the advanced economy need to be encouraged to align changes in the order of society with MSME activities in adapting to the e-digitalization period. This research has major implications for the development of MSMEs because MSMEs have succeeded in implementing e-commers in their sales and this greatly affects economic growth. Through the support of other analyses, it will also be known that the factors that support economic growth are more complete. This condition is more specific to developing countries and Covid-19 Pandemi.

Keywords: E-Commerce; Digitization of MSMEs; Economic Growth; Bigdata

1. Introduction

The Covid-19 pandemic that occurred was a severe shock to the world, entering Indonesia [1]. This pandemic has not only had a negative impact on health, but also for the economy[2][3]. The economic sector is an important sector that has been the focus of the Indonesian people. One of the provinces that has an impact on the economic sector is East Java Province. The sector that is particularly affected in East Java is the Micro, Small and Medium Enterprises (MSMEs) sector. MSMEs (Micro, Small and Medium Enterprises) are one of the drivers of the socio-economic sector. At the time of the economiccrisis of 1998, MSMEs are relatively able to survive compared to large companies. According to BPS data, after the economic crisis, the number of SMEs did not decrease. If this condition is not overcome. it will have a bad impact on economic growth in East Java. Therefore, it is necessary to find the right and correct solution for the condition of MSMEs to recover immediately[4]. Before the covid-19 pandemic, the number of MSMEs is increasing, even in 2012 it was able to absorb 85 million to 107 million workers. In that year, the number of entrepreneurs in Indonesia[5] was 56. 539. 560, where MSMEs 56. 534. 592 or about 99.99%.

This shows that MSMEs should be considered as one of the milestones in the Indonesian economy and employers. One solution is the utilization of e-commerce in MSMEs[6]. This research has an novelty because the analysis uses five different methodological approaches and instruments[7]. Through these 5 approaches, it is hoped that maximum results will be obtained.

This can be done because in the era of the Covid-19 pandemic[8][9], the shift in people's consumption patterns has changed from offline to online (e-commerce). If this happens quickly, it will affect economic growth. One of the advantages of e- commerce is that it has a very wide coverage area so that it allows economic interaction between regions in the era of the Covid-19 pandemic[10]. The presence of e-commerce is alsoa form of efforts to improve the quality of MSMEs through the digitization process. The development of SMEs is expected to help improve the welfare of local communities and accelerate national economic recovery. The pattern of the relationship between e-commerce and economic growth needs to be understood in order to adopt the right policies so that both can continue to grow. In addition, it is necessary to segment and optimize aspects that support the development of e-commerce and the digitization of MSMEs.

The social and economic impact of the Covid-19 pandemic has changed part of the order of people's lives. The existence of several social restrictions restricts traders, especially MSMEs (Micro, Small and Medium Enterprises) from direct sales [24]. Concerns aboutthe spread of Covid-19 have made people prefer to shop online. Economic digitalization is increasingly driving the transformation of people's shopping systems from offline to online.Therefore, e-commerce activities in Indonesia continue to increase during the Covid-19 pandemic. The growth of online/e-commerce transactions in the first semester of 2020 wasmarked by an increase in the value of e-money transactions in June 2020 which increased from 17.31% in May 2020 to 25.94% (year-on-year) (Bank Indonesia, 2020).

Based on the data obtained due to the impact of covid-19 on economic growth is important data that can be used by all people who need it. According to the journal "The Diffusion of Big Data and Its Impact on Official Statisticsand Statistical Agencies" big data can be a supporter of fundamental or classical statistics for mapping existing changes and identification in real- time. One of the sources of big data is social media. In Indonesia, online data sources from social media continue to grow along with the increase in internet users. Indonesia will have 202.6 million internet users by 2021, according to we are [11]"Numbers 2021: Indonesia". As many as 63.6% of internet users aged 16 to 64 years(production age) are internet users. I am very active on twitter. Most active Twitter users can be used as an online data source to take reactions and public opinion to government programs related to the digital economy.

The great development of technology indices in the current era and the disruption of technology bringing challenges Industry 5.0 is a challenge and opportunity that MSMEs must face in Indonesia ensuring the Competitiveness of SMEs Competing in the wider market The broad and significant impact of technology and uncertainty in the business environment[12]. So That Learning Opportunities Are Important the development of technology transfer models, Innovation and good knowledgecoming from the Government and academia will Encourage small businesses[13], medium and micro to grow and develop Indonesia into changing times and the current economy is very uncertain.

According to data from the Central Statistics Agency, the number of MSMEs in East Java is approximately 9,782,262 businesses. The number of e-commerce businesses in Indonesia

was around 15.08 percent in 2019 and grew to 90.18 percent in 2020. East Java Province has a percentage of e-commerce businesses that are always above the national percentage, which is 19.92% in 2019 and 90.31 percent in 2020. Other provinces in Java Island have a fairly large percentage of e-commerce businesses, in 2019 East Java Province had a fairly good economic growth of 5.52%, while in 2020 despite experiencing a recession (-2.39 percent) but the recession rate in East Java tended to be lower than other provinces in Java Island. This is relatively interesting to study, to see the impact of e-commerce business on economic growth in Java Island, segmenting and optimizing aspects that support the development of e-commerce in the focus of Java Island, especially East Java Province with the use of Bigdata to strengthen MSMEs in the batik industry sector.

The purpose of this study is to determine the relationship between e-commerce in MSMEs and economic growth with a multiple analysis approach. Through the results of this study, it is hoped that MSMEs[14] will be able to overcome the problems currently being faced. Research contributions are very important for MSMEs and the government because for MSMEs solutions will be obtained to overcome this, while for the government, it will be used to provide solution policies to the problems of the Covid-19 pandemic.

2.Theoretical background

2.1 Growth

According to BPS East Java (2021), the economic condition of East Java in terms of the value of GRDP on the basis of prices in force in 2020 has decreased when compared to 2019, which was IDR 46.33 trillion. Meanwhile, the value of East Java's GRDP on the basis of constant prices in 2010 in 2020 also decreased by RP 39.35 trillion from last year. The decline was caused by the COVID-19 pandemic that occurred since the beginning of 2020 which had an impact on the decline in business field production[15]. East Java has three main business fields that dominate the economic structure, namely the processing industry; large and retail trade, repair of cars and motorcycles; and agriculture, forestry, and fisheries. The business field contributed up to 60.51% in 2020. Gross Regional Domestic Product (GRDP) is the amount of added value generated by all business units in a given area, or is the sum of the final(net) value of goods and services produced by all economic units. The growth rate of Gross Regional Domestic Product (Economic growth) is obtained from the calculation of GRDP on the basis of constant prices. The growth rate is calculated by subtracting the GRDP value in the nth year against the value in the previous year n- 1, divided by the value in the n- 1st year, then multiplied by 100 percent. The growth rate points to the aggregate development of revenue from a given time against the previous time. The economic condition of East Java can also be seen in terms of economic growth. The economy of East Java during the period 2016-2019 experienced growth at a rate of 5%, but the economy of East Java in 2020 contracted to 2.39% as shown in Figure 1. This is because most business fields have experienced negative growth due to the COVID-19 pandemic.

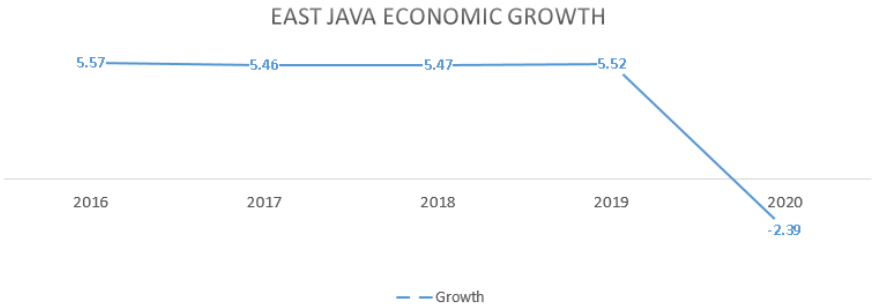

Figure 1. East Java Economic Growth in 2016-2020
Source: BPS Jawa Timur (2021)

Overall, the COVID-19 pandemic has disrupted economic activity in Indonesia. People who are unable to work due to this pandemic have experienced a decrease in income and a decrease in purchasing power[16]. In addition, sales in the industrial sector fell sharply, causing companies to be laid off because they could not afford to pay wages. The worst impact of COVID-19 in Indonesia, especially in the industrial sector, is that several large companies have gone out of business.

2.2 E-commerce Ventures

The rapid development of e-commerce can be an alternative for the community as a means of meeting needs during a pandemic. Various studies have illustrated the shift in consumer behavior in utilizing e-commerce to adapt to pandemic constraints such as health restrictions and activities imposed by the Government. According to Presidential Regulation Number 74 of 2017 and Government Regulation Number 80 of 2019 concerning Electronic Commerce, it is an electronic-based transaction system. E-commerce is a business that uses the internet to receive orders or sell goods/services . E-commerce refers to the application of modern computers, modern information and communication technology networks, and the application of electronic information technologies in the field of commerce. New forms of trade, commercial activities or non-profit services [18]. A report released by Insight Center and Kredivo (2021) shows that there is an increase in the number of consumer e-commerce users in 2020, namely 9 percent of consumers come from East Java. The value of online shopping transactions is also increasing, although it is uneven in some areas. Increased consumer e- commerce users. This is also supported by the increasing proportion of internet users to shop in East Java.

2.3 Bigdata for MSMEs

Big data technology can not only be used by large companies, but also by small and medium-sized enterprises (MSMEs)[17] and public organizations. Although big data technologies are considered complex and expensive, but small companies can also take advantage of big data as long as they know exactly what the purpose of the business is, thus facilitating the process of identifying the necessary data and obtaining greater benefits from investment. The implementation of Big Data strengthens the competitiveness of SMEs in Indonesia, challenges and opportunities for future development. The low digital literacy, technology transfer, diffusion and knowledge transfer make it a challenge that must be faced for the development of technology-based MSMEs in Indonesia. But this is not without the opportunity to remember millennials nowadays like technology-based startup business models. So that technology transfer opportunities are open. The benefits of big data that have been felt, especially for the business world, include knowing the public's response to the products issued through sentiment analysis in social media; helping the company make more precise and accurate decisions based on data; helps to improve the company's image in the eyes of customers; for business planning by knowing customer behavior, such as in telecommunications and banking companies; as well as knowing market trends and consumer desires According to [18] the benefits thatcan be used by companies by implementing big data are as follows, (1) social data analysis, (2) Historical data analysis, and (3) predictive analysis.

2.4 Research Framework

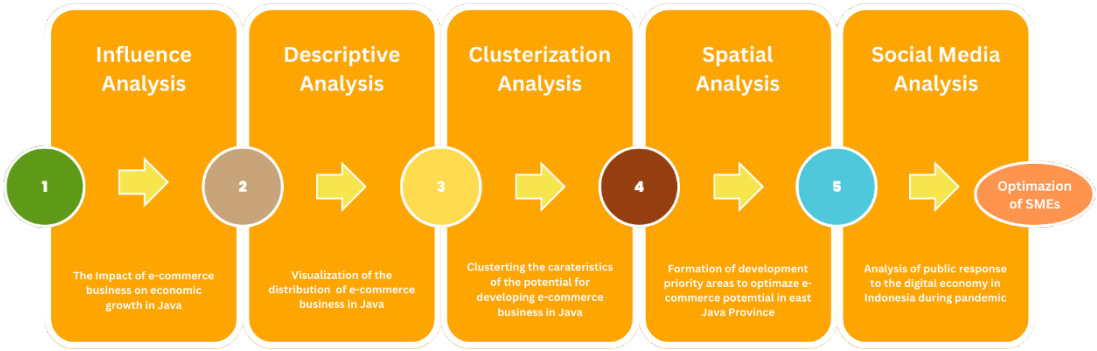

Figure 2 Research Framework

3. Search and Review Methodology

3.1 Research Variables and Data

In Table 1 below, it will be described about the variables and data used in the study. Variables will be described based on concepts and definitions as well as data sources for those variables[19]. The data used is secondary data consisting of a combination of classical statistical data and big data.

Table 1: Research Variables and Data

Variable	Concept and Definition
Economicgrowth	Comparison of the gross domestic product (GDP) of a region of the actual period to the GDP of the previous period. According to the Center Statistic Agency (BPS), GRDP is the amount of added value generated by all business units (economic sectors) in an area regardless of whether the production factors come from within the region or from outside the region. The GRDP value is presented in two types, namely GRDP on the basis of the prevailing price (ADHB) and GRDP on the basis of constant price (ADHK).
Percentage ofeffort *e-commerce*	Percentage of businesses that use the internet to take orders or sellgoods/services.
Human Dismantling Index(HDI)	An index that explains how residents can access development outcomes in earning income, health, education, and so on. HDI is formed by 3 (three) basic dimension, namely longevity and healthy living, knowledge, and a decent standard of living.
Percentage of villages with4G/LTE networks	Internet packet data system network with a certain data transfer rate. 4G signal is a type of 4th generation network from GPRS also called LTE (Long Term Evolution) which has a speed of up to 100 Mbps.
Percentage of villages with Base Transceiver Station(BTS) located	A BTS is a device that functions as a sender and receiver (transceiver) of cellular communication signals. BTS is marked by a tower / tower equipped with an antenna as a transceiver device. The public often refers to it as a cell phone tower.
Tweets	Tweet on social media Twitter with *the keyword* 'digital economy'.

Source: BPS

3.2 Research Analysis Techniques

The analysis used in this study consisted of five techniques. Such analysis techniques are as follows:

1. Influence Analysis (Regression)

Linear regression is an equation model that explains the relationship of non-free/bound variables (Y) with free/free variables. Regression can be used to find out the direction of the relationship or the influence of a free variable on a non-free variable. The relationship between one variable and one or more free variables can be expressed in linear regression models (Draper and Smith, 1992). In general the relationship can be expressed as the following equation:

$$Y = b0 + b1\,X1 + f \quad (1)$$

Where Y is a bound variable, X is a free variable, b0, b1 is the suspected parameter, and f is the regression error value.

The liner regression in this study will be used to test the effect of the percentage of e-commerce businesses on economic growth in a region. The locus used is the province of East Java. The period used was 2019 or one year before the pandemic occurred. The selection of such periods is used to obtain the result of pure influence when other conditions are normal. (ceteris paribus)

2. Descriptive Analysis (Thematic Map)

Descriptive analysis is carried out using thematic maps and bar charts[20]. Thematic maps are ableto present and visualize data more interestingly and more easily understood. In this study, thematic maps and bar charts will be used for looking at the distribution of the percentage of e-commerce business achievements between provinces in Indonesia in 2020 and 2021

3. Cluster Analysis (Geographically Weighted K-Means)

Cluster analysis (swarm) is an analysis that aims to group objects into relatively homogeneous groups based on the set of variables studied[21]. Bunch analysis is able to group objects based on the same. The characteristics of objects in a bunch have a high degree of similarity, while the characteristics between objects in a group and other groups have a low degree of similarity. In this study, the cluster analysis method used was a non-Hier-Arki method, namely K-Means. The K-Means method is used to place objects in a specific bunch based on the nearest mean. The K-Means method uses spatial position weighting, so that the position of the centroid does not change[22]. Thevariables that will be classified are two variables that affect the development of internet business,namely HDI, the existence of BTS and 4G LTE networks. The location used is a province on the island of Java, and the information period is 2020[23].

4. Spatial Analysis (Metropolitan Statistical Area)

Spatial analysis begins with testing spatial autocorrelation. Spatial autocorrelation can be described as whether there is a relationship between location and the similarity of features, ie. H. Conditions that are located close together have a greater impact compared to areas that are located far apart[12]. Spatial autocorrelation validation was carried out in two stages, namely worldwide spatial autocorrelation through the Moran worldwide index and local spatial autocorrelation through Moran scatterplot and local spatial correlation indicators[1]. The worldwide spatial autocorrelation of the Moran worldwide index is an autocorrelation that describes the relationship that usually occurs between the observed regions[24]. A positive I value indicates the occurrence of a rooted zone grouping with the same attribute, while a negative I value indicates a grouping of zones with different attributes. An I value close to zero indicates an unnatural relationship. Local spatial autocorrelation is used to measure extra-spatial relationships in more detail[25], such as B. Spatial grouping detects trends (spatial anomalies) in different regions of the surrounding region[26]. Moran distribution plots can be used to understand the stability of local spatial patterns, and local spatial association indicators can test the importance of spatial patterns[1]. In general, the concept of Metropolitan Factual Region (MSA) is an area consisting of the core area and the surrounding area, which unfortunately is very integrated with the core region economy[20]. In its development, MSA can be used to form priority areas as a basis for allocating funds, designing policies and implementing various government programs[20]. The MSA will be formed from lisa's bunch map to identify priority regions for development planning. On this basis, this study modified the MSA based on related research, namely, identifying priority areas for human development (HDI) and technology network development (4G LTE) in East Java. During the information period in 2020, the locations used were 38 regencies/cities in Tengan Province, East Java

5. Social Media Bigdata Analysis (Text Mining)

Text mining is defined as information mining compositions in which a user interacts with a set of documents using analytical tools that are components of information mining[11]. Tweets about Wear In this study used public gaps from various key customized structures of the advanced economy in Indonesia during the Coronavirus pandemic. The information used is an Indonesian tweet. Information retrieval is done using a Python application and the scratching method of the "Twint" crawler. Tweet information was collected from January 1, 2020 to June 30, 2021. The next stage after collecting information is sifting and text preprocessing using python. Twitter information usually has a lot of clamors so it is necessary to do text filtering and preprocessing. Filtering is done on rejecting information to get the information that is actually related to the customized structure of the key. Text preprocessing is performed to correct

information that is inappropriate and not ready for analysis, through case-collapsing, URL removal, punctuation removal, stop word, and so on. Information ready for analysis is 9,605 tweets Sentiment analysis or opinion mining is a branch of SMS research that aims to determine the public's perception or objectivity towards a topic, event, or problem. A lexicon-based approach is a sentiment analysis technique that involves calculating the polarity of sentiment by using semantic orientations or sentences in text or documents [8]. The analysismethod of this study uses sentiment analysis and descriptive analysis with R in the form of line graph visualization for the entire descriptive tweet information, circle charts and line outlines for customized organization keys for sentiment analysis visualization, and conjunctions for visualization. from the expectations of the government contained in the tweet information.

The steps taken in carrying out analysis and visualization are as follows.

1) Classification of sentiment data using a manual weighting method for each word based on a dictionary of lexicon [25]. So that it can be identified as negative, neutral, and positive sentiment. If the result weighs less than one (-- deleted HTML --) 0) then the tweet is categorized as positive.
2) After classifying sentiment data, each tweet is continued with visualization using pie charts and line charts/trend data so that it can be known how digital economy sentiment is developing.
3) Next, process tokenization and merging of words in tweets. Tokenization is done to separate words, then the second to third words are recombined. The first words that are still separate are chosen for words that have a special meaning of expectation, so that the data can be classified into the expectations of the digital economy.
4) The final step is to visualize the sentence of expectation with a connecting word.

4.Results and Discussion

4.1 Effect Analysis (Regression) Simultaneous Test (F)

The hypotheses used to test are H0: The unfit model (e-commerce ventures do not affect economic growth) and H1: The fit model (e-commerce businesses hinder economic growth). Based on the results of the F test, a p-value of 0.072 was obtained which is smaller than the alphavalue of 7.5%, meaning Reject H0. So, the results of the simultaneous test F showed that the model fit at a value of α = 7.5%. Parameter Estimation Results. The results of the estimation usinglinear regression showed that independent variables and constants were significant at α = 7.5%. The variable percentage of e-commerce businesses significant affects the economic growth of provinces in Java Island.

Table 2: Results of Estimating Parameters of the Economic Growth Equation Model

Variabel	Koefisien	Std. error	T-Statistics	p-value
Constanta	3,7484	0,7852	4,7734	0,0088
Usaha *e-commerce*	0,0945	0,0389	2,4257	0,0723
R^2 = 0,5953				

Source: Data Processing Summary

Based on the yield in Table 2, it can be concluded that the percentage of e-commerce businesseshas a positive effect on economic growth, a change in the minimum percentage of internet business businesses by 1 percent will cause the economy to grow by 0.0945 percent. This is in accordance with the results of research by Liu (2013) and Anvari and Nourozi (2016) who claim that web-based business has a positive effect on economic growth. The value of the coefficient of determination (R2) of 59.53% means that the variation in economic growth in 2019, around 59.53 percent, can be explained by the number of internet business businesses, while the variation of 40.47 percent is explained by other variables that are not included in the

model. The equation model of the percentage of internet business ventures formed:

$$\widehat{Ecommerce}_t = 3,7484 + 0,0945 pertumbuhan\ ekonomi_t \quad (2)$$

4.2 Residual Assumption Testing

The model of the system of linear regression equations formed needs to be tested for residual assumptions of modeling results. The tests include normality and homogeneity tests. Based on the results in Table 3, the assumptions of normality and homochedasticity have been met at the value of α = 7.5%.

Table 3: Normality and Homogeneity Test

Persamaan	Uji Normalitas (Jarque-Bera)		Uji Homogenitas (Breusch-Pagan)	
	Statistik	p-value	Statistik	p-value
Regresi linear	0.7181	0.6983	0.0097	0.9217

Source: Data Processing Summary

Descriptive Analysis (Thematic Map)

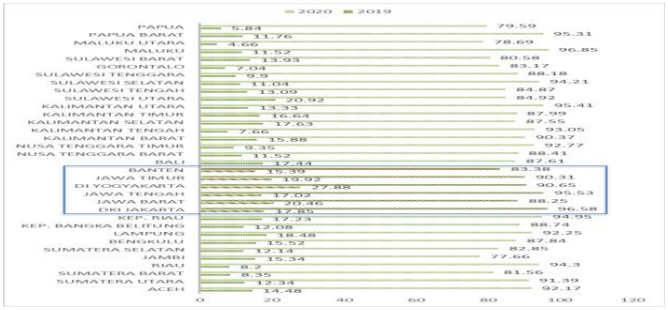

Figure 2: Percentage of e-commerce Businesses per Province for the 2019 and 2020 Period

In 2019, the percentage of online business in Indonesia nationally was 15.08 percent, while in 2020 it grew to 90.18 percent. In general, all provinces also experienced an increase in online business in 2020, in line with changes in socioeconomic patterns in society due to the Coronavirus pandemic. If observed in the scope of Java Island, in 2019 all provinces in Java Island havea higher percentage of online business than the percentage nationally. With a value of 19.92 percent, East Java province became the third highest province in Java Island, after DI Yogyakarta and West Java. In 2020, the six provinces experienced an increase in the number of online businesses, but no longer all provinces in Java Island have a higher percentage of internet business businesses than the national percentage.

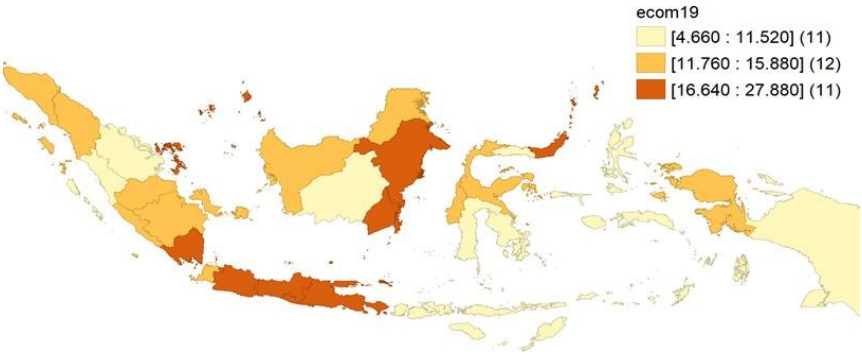

Figure 3: E-commerce Percentage Distribution Map for the 2019 Period

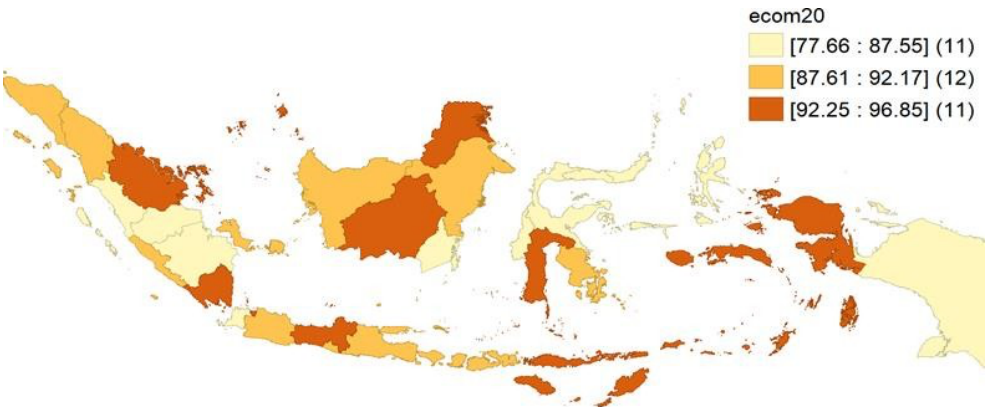

Figure 4: E-commerce Percentage Distribution Map for the 2020 Period

Based on Figure 3, it can be seen that if divided into 3 quantiles, then in 2019 almost all provinces on the island of Java were included in the third quantile with a high value of online business business compared to the other two quantiles. However, based on Figure 4, it can be seen thatin 2020, only Central Java and DKI Jakarta Provinces were part of quantile 3. This indicates that the compositions of the computerized adaptation of MSME in the form of online businessin Java Island are still not as fast as in other provinces. Optimizing the digitization of MSMEs and the development of internet business can be done by optimizing the aspects that affect it. Therefore, in the third stage, a cluster analysis will be carried out to see the segmentation of aspect variables related to sanctum gan online business development.

4.3 Cluster Analysis (Geographically Weighted K-Means)

By Research HDI shows that people competence and human competence are positively related to adoption, especially technology such as those used in online business. The next variable that has a porous effect on the percentage of online business businesses is the percentage of villages that have BTS. Another variable that affects the percentage of online business is the percentage of villages with 4G web networks. This variable also positively affects the percentage of e-commerce businesses. These two variables are in line with the results of research by Waseem et al. (2018) that communication and information technology infrastructure has a large and positive contribution in building an online business. Aspects that affect the optimization of online business development, namely HDI, bts existence, and 4G networks in 2020 in Java Island will be grouped with cluster analysis. Here are the results of the cluster analysis of the K-Means method with location geographical as weight.

Figure 5. Formed Cluster

Based on Figure 5 it can be seen that three clusters are formed. Cluster one consists of 3 provinces, cluster two with 2 provinces, and cluster three with 1 province. East Java Province is incorporated into cluster one with Central Java and DI Yogyakarta

Table 5: Characteristics of K-Means Cluster Results

Klaster	IPM	Keberadaan BTS	Jaringan 4G LTE
(1)	(2)	(3)	(4)
1	74,52	58,23	83,54
2	76,61	79,46	87,07
3	72,09	71,46	87,71

Source: Data Processing Summary

Table 5 contains the characteristics of each cluster. It can be noticed that cluster two has good withering characteristics. Cluster one has an HDI value that is as high as compared to the other two clusters, but has the lowest BTS presence and 4G LTE network. Next, further analysis will be carried out regarding the detailed conditions of the HDI and the quality of the inter-district network or cities in East Java Province so that appropriate wisdom can be taken for the optimization of advanced MSMEs in East Java.

4.4 Spatial Analysis (Metropolitan Statistical Area) Human Development Index

Spatial analysis begins with checking global autocorrelation to determine whether there is a spatial influence on HDI in East Java Province. Based on the results of data processing (Figure 5), a Global Moran's I value of 0.4538 was obtained, so it can be concluded that spatial autocorrelation occurred in HDI in East Java Province. This means that there is a mutual relationship between regencies/cities in East Java Province based on the HDI value.

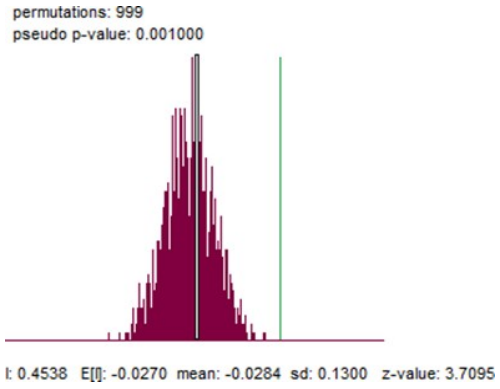

Figure 6: Global Autocorrelation of HDI of Regencies/Cities in East Java, 2020

The value of the Moran's I Global Statistics obtained was significant at the test level (α = 7.5%), with a pseudo p-value = 0.001 at permutation 999 times. A positive Global Moran's I value indicates the occurrence of positive spatial autocorrelation in the form of location looming that has similar HDI characteristic values. Furthermore, local spatial autocorrelation checks were carried out to determine the spatial pattern of HDI in East Java Province and identify which districts/cities were spatially grouped.

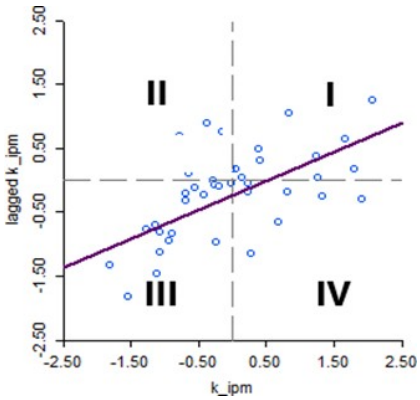

Figure 7: Moran's Scatterplot for HDI in East Java Province, 2020

In Figure 7, the Moran;s scatterplot method is used to determine the pattern of spatially pumping HDI. Moran's scatterplot shows that most of the regencies/cities in East Java Province are in quadrants I and III. This indicates the existence of positive spatial autocorrelation (in line with theGlobal Moran's I values obtained at the previous stage of the analysis). There are several districts/cities that fallow in quadrants II and IV (spatial outliers). Based on the classification of areas according to the quadrant in Figure 6, the process of identifying hotspots and coldspots is then carried out using the LISA cluster map.

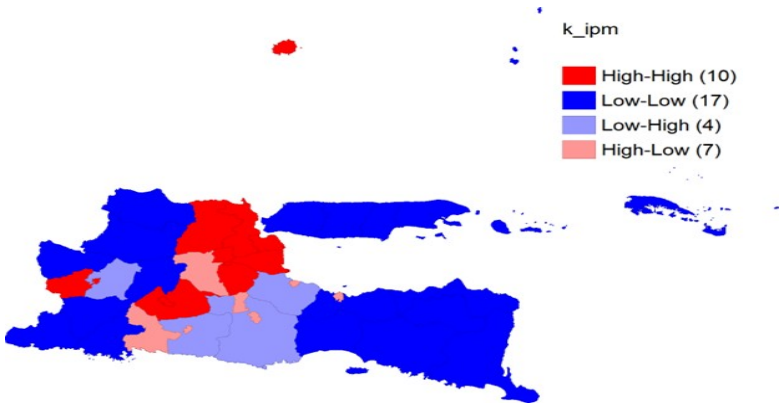

Figure 8: LISA Cluster Map for HDI in East Java Province, 2020

Based on Figure 8, 10 hotspots (High-High) were identified, namely Sura-baya City, Gresik Regency, Lamongan Regency, Mojokerto City, Mojokerto Regency, Sidoarjo Regency, Kediri City, Kediri Regency, Madiun City, and Magetan Regency. This means that there is a grouping of regencies/cities with high HDI values around the ten regions. This reflects that development has been successfully carried out in the region and the surrounding area.
It can also be seen that there are 17 coldspot points (Low-Low) that tend to be ter- bisa on the left and right sides of East Java Province. In addition, there are 7 areas that are included in the spatial outliers (High-Low) category, namely Jombang Regency, Batu City, Malang City, Pasuruan City, Probolinggo City, Tulungagung Regency, and Blitar City. The High-Low criterion shows that the seven regions have relatively high HDI compared to the surrounding regions.

Based on the identified hotspots, coldspots, and spatial outliers, a Metropolitan Statistical Area (MSA) analysis was then carried out. The MSA is adapted from OMB. The MSA will be formed based on the LISA cluster map to determine the priority areas of human development planning.
Regencies/cities that are a priority for human development are districts/cities with low HDI achievements. Central county/core is a district/city den- gan achieving a relatively higher HDI than the surrounding area. It is identified based on the High-Low category on the LISA cluster folder. Outlying counties are counties around the core that have high socioeconomic ties to the core.

In this study, 2 MSA were formed which are priori- bag areas of human development. First, on the left side, MSA Jombang Tulungagung is formed. The MSA consists of Jombang Regency and Tulungagung Regency as central coun- ty/core, and 7 other districts as outlying counties. Jombang and Tulun-gagung regencies have HDI values that continue to grow continuously from year to year, until in 2020 it reached the values of 73 (Tulungagung) and 72.97 (Jombang). The increase in HDI shows the success of human development in both districts, so that it cannot be used as a pilot material for the 7 outlying districts. The second MSA, on the right side, is the Pasuruan MSA. The MSA consists of Pasuru City as a central county/core, and 6 other districts as an outlying county. Human development in the Pasuruan MSA can be aligned with the results of the RKPD (Regional Government Plan) of Pasuruan Regency 2020 which will focus on human resource-based development with the theme 'Quality Human Development and Public Servicesto Improve Regional Competitiveness' (Pasuruan Regency Government, 2019).

4.5. G LTE Network

Furthermore, spatial analysis of the 4G LTE network through checking the global autoco- relationship to determine whether there is a spatial influence on the 4G network in East Java Province. Based on the results of data processing (Figure 8), a Global Moran's I value of 0.2292 was obtained, meaning that spatial autocorrelation or interrelationships between regions of districts/cities in East Java Province were obtained based on the availability of 4G networks.

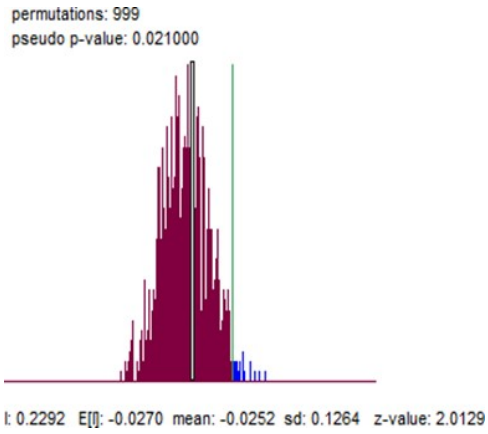

Figure 9: Global Autocorrelation of 4G Network of Regencies/Cities in East Java, 2020

The value of the Moran's I Global Statistics obtained was significant at the test level (α = 7.5%), with a pseudo p-value = 0.021 at 999 permutations. A positive Global Mo- ran's I value indicates the occurrence of positive spatial autocorrelation i.e. the grouping of locations with similar 4G network characteristics. Furthermore, local spatial autocorrelation checks were carried out to determine the spatial pattern of the 4G network in East Java Province and identify which districts/cities are spatially grouped

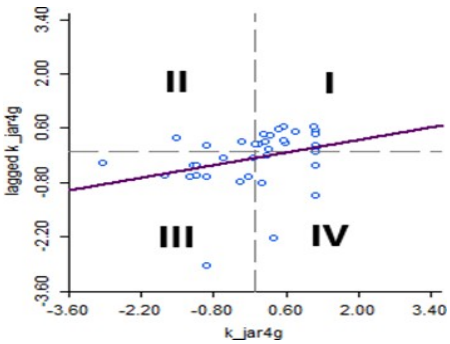

Figure 10: Moran's Scatterplot for 4G Network in East Java Province, 2020

In Figure 10, Moran's scatterplot method is used to determine the spatial grouping pattern of 4G networks. Moran's scatterplot shows that most of the districts/cities in East Java Province are in quadrants I and III. It shows the presence of positive spatial autocorrelation (in line with the Global Mo- ran's I values obtained in the previous stage of the analysis). There are several districts/cities that are in quadrants II and IV (spatial outliers). Based on the classification of regions according to the quadrant in Figure 9, the process of identifying hotspots and cold-spots is then carried out using the LISA cluster map

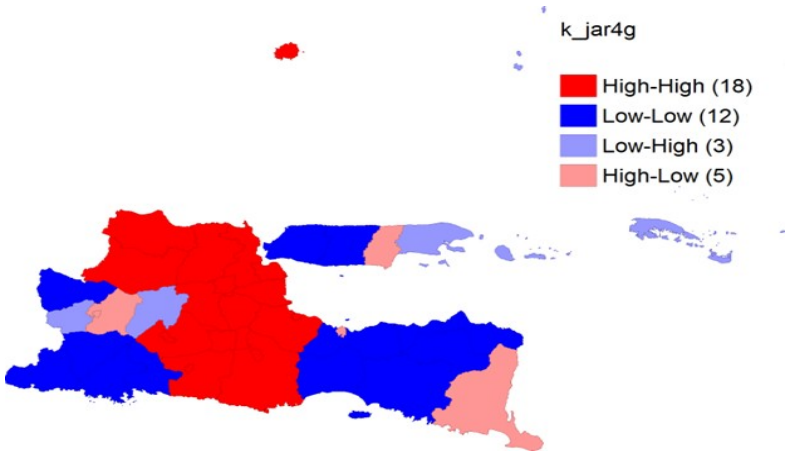

Figure 11: LISA Cluster Map for 4G Network in East Java Province, 2020

Based on Figure 11, 18 hotspots (High-High) were identified that tend to cluster in the central part of East Java Province. This means that there is a grouping of districts/cities with high 4G network values in around eighteen of these areas. This reflects that network infrastructure development has been successfully carried out in the region and the surrounding area (the central region of East Java Province).
It can also be seen that there are 12 cold spot points (Low-Low) that tend to be ter- bisa on the lower left side and the right side of East Java Province. In addition, there are 5 areas that are included in the spatial outliers (High-Low) category, namely Banyuwangi Regency, Pa-suruan City,Madiun Regency, Madiun City, and Pamekasan Regency. The High-Low criterion shows that all five regions have relatively better 4G networks compared to the surrounding regions.

Based on the identified hotspots, coldspots, and spatial outliers, a Metropolitan Statistical Area (MSA) [27]analysis was then carried out. The MSA is adapted from OMB. The MSA will be formed based on the LISA cluster map to determine priority areas for planning internet network infrastructure development. Regencies/cities that are priorities for human development are districts/cities with low 4G network availability. Central county/core is a district- en/city with a relatively higher 4G network than the surrounding area. It is identified based on the High-Low category on the LISA cluster map. Outlying counties are counties around the core that have high socioeconomic ties to the core[28].

In this study, 3 MSA can be formed which are areas prior to infrastructure development. First, on the lower left side, MSA Madiun is formed. The MSA consists of Madiun Regency and Madiun City as the central county/core, and the other 5 districts as outlying counties. The second MSA, on the upper right side, is MSA Pamekasan. The MSA consists of Pamekasan Regency as the central county/core, and 3 other districts as outlying counties. The third MSA, on the upper right side, is MSA Banyuwangi Probolinggo. The MSA consists of Banyuwangi Regency and Probolinggo City as the central district/core, and 5 other districts as outlying counties. Equitable distribution of internet network infrastructure can be focused on MSA-MSA which is formed by using the core as a development center. Comparative studies with districts/cities that have been categorized as High-High also needs to be done to speed up the process of developing the network and accelerate the procedures for disseminating the adaptation of digital technology to its use in various villages. The process of digitization to various villages can help optimize the development of MSMEs and the economic structure in the village.

Social Media Big Data Analysis (Text Mining)

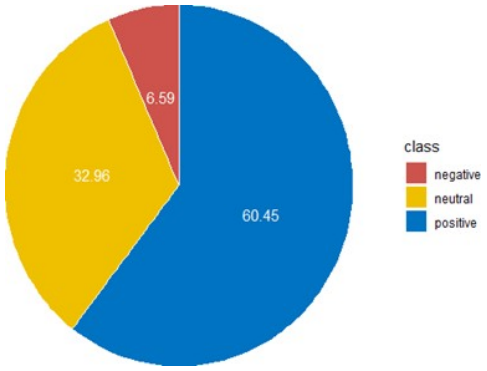

Figure 12: Percentage of Public Responses related to the Digital Economy (January 2020 – June 2021)

Figure 12 shows the percentage of public opinion about the digital economy. Based on the image, the response was dominated by tweets with positive and neutral opinions, namely 60.45 percentand 32.96 percent. It can be concluded that most of the people are recive and support the digital economy. The amount of positive public opinion suggests that the digital economy deserves to be continued by the government. However, there are still 6.59 percent of public opinion that criticizes or disagrees with the digital economy.

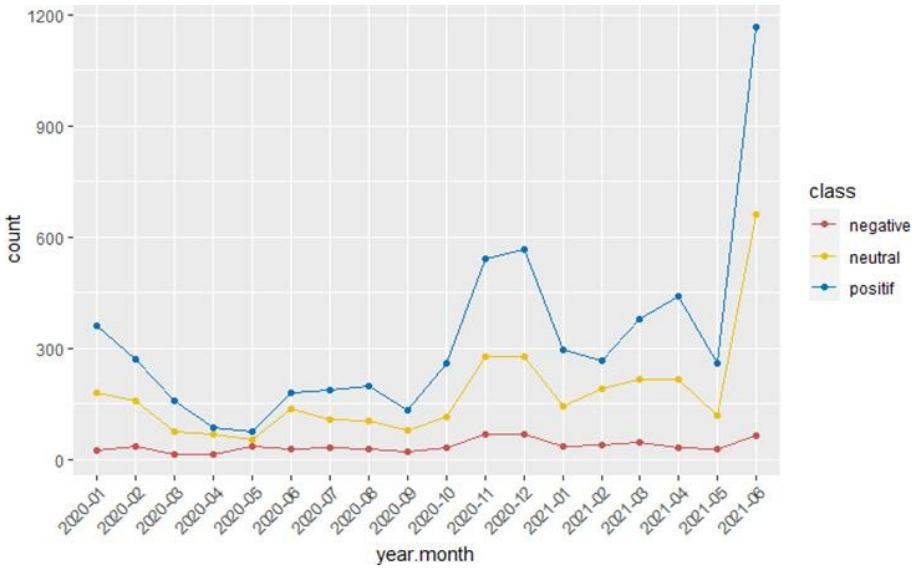

Figure 13: Public Response Trends related to the Digital Economy (January 2020 – June 2021)

Figure 13 shows that the trend of digital economy sentiment from the beginning of 2020 to June 2021 was dominated by tweets with positive opinions and neutral opinions, while negative opinions were far below it. This is consistent with Figure 11 regarding the percentage of cents on the digital economy which is dominated by positive and neutral sentiment. The positive responseof the digital economy has a trend that continues to increase over time, which shows that peopleare increasingly aware of the importance of the digital economy.

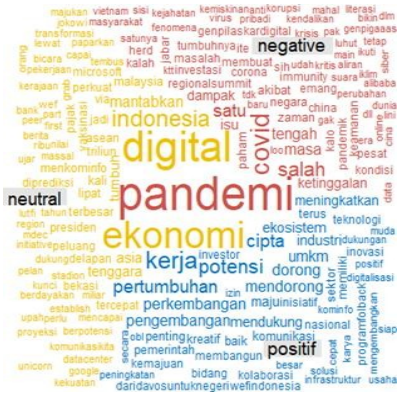

Figure 14: Wordcloud Digital Economy

Figure 14 is a digital economy word cloud showing that public opinion about the digital economy is very diverse. Neutral opinion is dominated by a description of the activity and some public expectations. Negative opinions are highlighted by the words pandemic which shows that the digital economy needs to be optimized to reduce the impact of poverty during the pandemic. The following are examples of some tweets for each opinion. An example of a tweet with a neutral opinion sentiment is "China Asean Expo paves the way for cooperation in the digital economy sector". Positive opinions are dominated by tweets from people who feel help or from program implementers who have implemented the program. An example of a positif tweet is "The digital economy is not only good for internal SMEs but also overall economic growth as the flow of goods outside the city increases, the logistics and industrial sectors also have a cool positive impact". Furthermore, the following are examples of tweets with negative opinions on the digital economy, namely "Not much has been revealed about the negative impact of the carbon footprint of the digital economy, modern technology, tablets, smartphones, televisions, to electric cars, depending on rare materials from the earth" and "The digital economy in Indonesia has many challenges, internet penetration is still lagging behind other ASEAN countries, besides cybersecurity is still poor.

Imported products flooded the average low internet speed and only 1 million SMEs out of millions went digital" [29]. Based on the example of the negative tweet, it can be seen that negative opinions indicate a problem in the form of public criticism of the digital economy[30]. The government needs to optimize the use of the internet and empower MSMEs to go digital, as well as strive for innovation in the development of sustainable digital products.

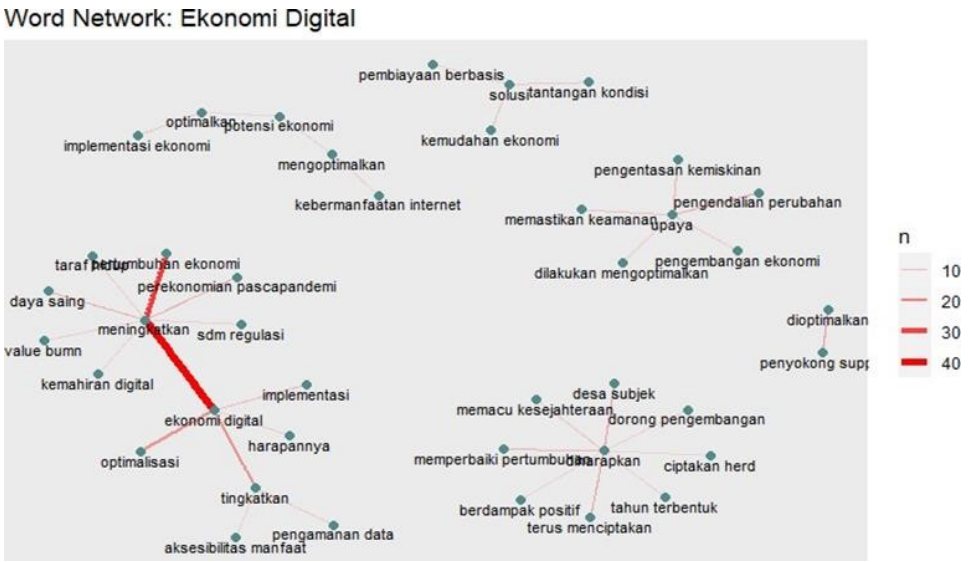

Figure 15: Digital Economy Linkword

Figure 15 is a visualization of some of the expectations that are often issued by the public towards the digital economy in the form of linkword visualizations. According to Figure 15, the hope that often comes out of public tweets related to the digital economy is that it can be a solution to improve economic growth, optimize demand and supply, and improve welfare. The community also provides input in this program, namely the government must optimize the digital economy to maintain the quality of the workforce, data security, internet benefits, and economic development. All optimalization of the digital economy that is aligned with the progress of MSMEs can absorb labor, drive the economy, and change the pandemic situation to accelerate the national economic recovery[5].

5. Conclusion

The e-commerce business significantly affects the economic growth of the provinces in Java. The Covid-19 pandemic caused the number of e-commerce businesses in each province to increasein 2020. The improvement needs to be aligned with the quality of aspects that affect the development of e-commerce. However, based on cluster analysis, East Java Province is still classified as a cluster that has a HDI and the quality of the internet network is not good compared to other provinces in Java Island. The results of the Metropolitans Statistical Area (MSA) show that it is necessary to optimize the achievement of HDI and the quality of the internet network, especially in regencies/cities located on the left and right sides of East Java Province. Based on social media analysis through text mining in Indonesia in the pandemic era (January 2020 to June 2021), the digital economy received a positive response of up to 60.45 percent and a neutral response of 33.16 percent. The large number of positive responses shows that the digital economy has managed to get support from the community. People hope that the digital economy can be a solution to improve economic growth, optimize demand and supply, empowerMSMEs go digital, and improve welfare. The development of human resource quality, network infrastructure development, and increasing digital economic literacy are expected to be key in empowering MSMEs and optimizing e-commerce in East Java Province in order to overcome the impact of the pandemic while accelerating economic recovery.

In terms of regulations, government support for e-commerce has been carried out through the issuance of Presidential Regulation No. 74 of 2017 concerning the road map of the electronic- based national trade system or e-commerce and then strengthened by the issuance of Government Regulation Number 80 of 2019. The effect of e-commerce business on economic growth is the basis for the need to optimize the development of e-commerce. The optimization process is increasingly needed because it is related to the lack of quality of e-commerce aspects in East Java Province based on hdi values and internet network availability. In an effort to optimize HDI and internet networks in East Java Province, the government needs to adapt the penta-helix mechanism. Penta-helix is an economic framework that combines coordination between government, academia, the business sector, society, and the media at the same time. The government plays the role of a regulator. Academics will become research designers as well as literacy agents for human resource empowerment programs in MSMEs and digitize the use of the internet in the community. The business sector acts as a supporter of the infrastructure development process and facilitator for the successful benchmark of MSME empowerment. The public and the media act as object as well as dissemination of various activities to support every program and policy that is being carried out by the government in optimizing the empowerment of MSMEs through the development of human resources and quality internet networks.

In terms of the competence of human resources to support the improvement of e-commerce businesses in East Java, it is necessary to hold various trainings to increase the expectations of MSMEs, business actors or prospective e-commerce business actors regarding online selling techniques. This can be actualized by the existence of a mentoring program for online business actors in order to increase their online transactions, increase the scale of their business, or even encourage business actors to increase market reach to be international or go international (export market share). In terms of facilities, the percentage of villages with the existence of BTS can bridge user communication devices with the acquisition of 4G network quality. A quality internet network can guarantee the speed of data transfer to support the digitization process of MSMEs. In terms of public services, the government needs to pay more attention to the shortcomings of the digital economy based on negative sentiments from the public as an effort to increase public trust and digital alliteration to the public. The government can use social media

analysis with the text mining method as an alternative to evaluating government programs related to the digital economy. In the future, it is necessary to create an application or dashboard system to display real-time evaluation of government programs through social media analysis. The system is expected to speed up the evaluation process so that the taking of policies related to the empowerment of MSMEs through the digital economy can be carried out appropriately and responsively.

5. Limitations and future research

The limitation of this study is that the discussion is only limited to one country so that it cannot provide general conditions for MSMEs. This condition is more specific to developing countries. In addition, the detailed discussion is only focused on digital marketing, namely E-Commers, which can actually be discussed about the availability of capital etc. The results of the study are expected to be a reference for cases of MSMEs affected by the Covid-19 Pandemic in developed countries.

References

[1] L. Anselin, “Local Spatial Autocorrelation (1), Common Univariate Local Statistics.” 2018.
[2] R. D. Anvari and D. Norouzi, “The impact of e-commerce and R&D on economic development in some selected countries,” *Procedia-Social Behav. Sci.*, vol. 229, pp. 354–362, 2016.
[3] E. Rosillo-Díaz, F. J. Blanco-Encomienda, and E. Crespo-Almendros, “A cross-cultural analysis of perceived product quality, perceived risk and purchase intention in e-commerce platforms,” *J. Enterp. Inf. Manag.*, 2019.
[4] E. Halim and R. H. Karami, “Information systems, social media influencers and subjective norms impact to purchase intentions in e-commerce,” in *2020 International Conference on Information Management and Technology (ICIMTech)*, 2020, pp. 899–904.
[5] T. Tambunan, “Recent evidence of the development of micro, small and medium enterprises in Indonesia,” *J. Glob. Entrep. Res.*, vol. 9, no. 1, pp. 1–15, 2019.
[6] A. Wanto *et al.*, “Model of Artificial Neural Networks in Predictions of Corn Productivity in an Effort to Overcome Imports in Indonesia,” in *Journal of Physics: Conference Series*, 2019, vol. 1339, no. 1, p. 12057.
[7] P. Chiranjeevi, D. Teja Santosh, and B. Vishnuvardhan, “Survey on sentiment analysis methods for reputation evaluation,” in *Cognitive Informatics and Soft Computing*, Springer, 2019, pp. 53–66.
[8] N. Astari, “A Literature Review: Digital Marketing Trends in Indonesia During the COVID-19 Pandemic,” *CHANNEL J. Komun.*, vol. 9, no. 2, p. 125, 2021.
[9] A. J. R. Basu, A. Khatua, S. Ghosh, and A. Jana, “Harnessing twitter data for analyzing public reactions to transportation policies: Evidences from the odd-even policy in Delhi, India,” *Proc. East. asia Soc. Transp. Stud.*, 2017.
[10] S. Basuki and R. Anugrah, “Transaction Document Security Protection In The Form Of Image File, Jpg Or Tif Interbank Transfer Using Steganography And Cryptography,” *IAIC Trans. Sustain. Digit. Innov.*, vol. 1, no. 1, pp. 42–48, 2019.
[11] S. Dodd, M. Clarke, L. Becker, C. Mavergames, R. Fish, and P. R. Williamson, “A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery,” *J. Clin. Epidemiol.*, vol. 96, pp. 84–92, 2018.
[12] M. J. de Smith, *Statistical analysis handbook*. The Winchelsea Press, 2018.
[13] N. Zainuddin, A. Selamat, and R. Ibrahim, “Hybrid sentiment classification on twitter aspect-based sentiment analysis,” *Appl. Intell.*, vol. 48, no. 5, pp. 1218–1232, 2018.
[14] B. J. Kiplangat, A. Shisia, and I. C. Asienga, “Effects of human competencies in the adoption of e-commerce strategies among SMEs in Kenya,” 2015.
[15] B. B. Hapsoro, P. Palupiningdyah, and A. Slamet, “Peran digital marketing sebagai upaya peningkatan omset penjualan bagi klaster UMKM di Kota Semarang,” *J. Abdimas*, vol. 23, no. 2, pp. 117–120, 2019.
[16] N. Brachetta-Aporta and L. A. D’addona, “Applying multivariate methods to dental development sequences for subadults from archaeological samples,” *Int. J. Osteoarchaeol.*, vol. 30, no. 2, pp. 218–224, 2020.
[17] A. A. Mattjik, I. Sumertajaya, G. N. A. Wibawa, and A. F. Hadi, “Sidik peubah ganda dengan menggunakan SAS.” 2011.
[18] G. Liu, S. Fei, Z. Yan, C.-H. Wu, S.-B. Tsai, and J. Zhang, “An empirical study on response

to online customer reviews and E-commerce sales: from the mobile information system perspective," *Mob. Inf. Syst.*, vol. 2020, pp. 1–12, 2020.
[19] U. Rahardja, "Blockchain Education: as a Challenge in the Academic Digitalization of Higher Education," *IAIC Trans. Sustain. Digit. Innov.*, vol. 4, no. 1, pp. 62–69, 2022.
[20] R. Vought, "Revised delineations of metropolitan statistical areas, micropolitan statistical areas, and combined statistical areas, and guidance on uses of the delineations of these areas," *OMB Bull*, no. 20–01, 2020.
[21] J. Pirkis *et al.*, "Suicide trends in the early months of the COVID-19 pandemic: an interrupted time-series analysis of preliminary data from 21 countries," *The Lancet Psychiatry*, vol. 8, no. 7, pp. 579–588, 2021.
[22] H. Sularsih and A. Nasir, "Strategi UMKM dalam meningkatkan pendapatan dimasa pandemi COVID-19 guna mempertahankan kelangsungan usaha di era revolusi industri 4.0 (studi pada UMKM makanan dan minuman di Kabupaten Pasuruan)," *J. Paradig. Ekon.*, vol. 16, no. 4, pp. 763–772, 2021.
[23] P. RI., *Peraturan Pemerintah Nomor 80 Tahun 2019*. 2017.
[24] B. Zhang, Z. Jiang, Y. Chen, N. Cheng, U. Khan, and J. Deng, "Geochemical Association Rules of Elements Mined Using Clustered Events of Spatial Autocorrelation: A Case Study in the Chahanwusu River Area, Qinghai Province, China," *Appl. Sci.*, vol. 12, no. 4, p. 2247, 2022.
[25] S. A. Faaroek, A. S. Panjaitan, Z. Fauziah, and N. Septiani, "Design and Build Academic Website with Digital Certificate Storage Using Blockchain Technology," *IAIC Trans. Sustain. Digit. Innov.*, vol. 3, no. 2, pp. 175–184, 2022.
[26] S. Biduri and B. Proyogi, "Acceleration of performance recovery and competitiveness through non-banking financing in SMEs based on green economy: impact of Covid-19 pandemic," *J. Innov. Entrep.*, vol. 10, no. 1, pp. 1–10, 2021.
[27] A. Waseem, "Analysis of Factor Affecting e-Commerce Potential of any Coun-try using Multiple Regression."
[28] P. Bergquist, D. M. Konisky, and J. Kotcher, "Energy policy and public opinion: patterns, trends and future directions," *Prog. Energy*, vol. 2, no. 3, p. 32003, 2020.
[29] S. A. Sarkodie and I. Ozturk, "Investigating the environmental Kuznets curve hypothesis in Kenya: a multivariate analysis," *Renew. Sustain. Energy Rev.*, vol. 117, p. 109481, 2020.
[30] L. F. FG Assis *et al.*, "TerraBrasilis: a spatial data analytics infrastructure for large-scale thematic mapping," *ISPRS Int. J. Geo-Information*, vol. 8, no. 11, p. 513, 2019.