Journal of Economics, Business, and Accountancy Ventura Vol. 27, No.1, April - July 2024, pages 14-32

The Effect of E-Commerce on Gross Regional Domestic Product and Clustering of Its Characteristics by Utilizing Official Statistics and Big Data

Widhelia Echa Pramesthy, Putri Muthi'ah Dini, Muhammad Amirul Budiman, Zul Ahmad Esharja, Fitri Kartiasih*

Politeknik Statistika STIS, Jakarta Timur, DKI Jakarta, Indonesia

A R T I C L E I N F O

Article history:
Received: September 11, 2023
Revised: May 15, 2024;
May 30, 2024
Accepted: May 31, 2024

JEL Classification:
C32, O11, O40

DOI:
10.14414/jebav.v27i1.4136

Keywords:
GRDP, E-commerce, Cluster, Big data, Google trends

A B S T R A C T

The objective of this study is to examine the impact of e-commerce on Gross Regional Domestic Product (GRDP) and categorize GRDP according to e-commerce features in 17 provinces throughout Java, Sumatra, and Bali. The Human Development Index (HDI), which measures people's proficiency with technology—among which is e-commerce—is greatest on these three islands. This study makes use of big data in the form of Google Trends data as well as official statistics data obtained from the BPS-Statistics Indonesia. K-means clustering and regression analysis are the research methods employed. The study's findings demonstrate that the three islands' GDP growth rates are considerably accelerated by e-commerce companies. There were two ideal clusters identified: the low and high GRDP clusters. All of the provinces on the island of Java are included in the high cluster, with the exception of Central Java, which is distinguished by high transaction values and e-commerce-related keyword searches. Provinces in the low cluster are located on the Indonesian islands of Sumatra and Bali, and they are distinguished by the usage of Cash on Delivery (COD) and a limited number of employees in e-commerce businesses.

A B S T R A K

Penelitian ini bertujuan untuk menganalisis pengaruh e-commerce terhadap Produk Domestik Regional Bruto (PDRB) dan mengelompokkan PDRB berdasarkan karakteristik e-commerce yang mencirikannya dengan lokus 17 provinsi di Pulau Jawa, Sumatera, dan Bali. Ketiga pulau ini memiliki Indeks Pembangunan Manusia (IPM) tertinggi yang dapat menggambarkan suatu kecakapan masyarakat dalam bidang teknologi, salah satunya e-commerce. Studi ini menggunakan data official statistics bersumber dari Badan Pusat Statistik (BPS) dan big data berupa data Google Trends. Metode penelitian yang digunakan adalah analisis regresi dan K-Means Cluster. Hasil penelitian menunjukkan bahwa usaha e-commerce signifikan meningkatkan laju pertumbuhan PDRB di ketiga pulau tersebut. Ditemukan dua klaster optimal, yaitu klaster PDRB tinggi dan rendah. Provinsi pada klaster tinggi terdapat di Pulau Jawa, kecuali Jawa Tengah yang dicirikan oleh nilai transaksi dan pencarian keyword terkait e-commerce yang tinggi. Provinsi pada klaster rendah terdapat di Pulau Sumatera dan Bali yang dicirikan oleh jumlah tenaga kerja usaha e-commerce yang sedikit dan menggunakan metode pembayaran langsung (COD).

1. INTRODUCTION

Information systems are a topic that continues to attract worldwide attention. Starting from the 20th century, Internet use became increasingly widespread. At first, the Internet was used as a security system, but now it has developed to be used in various fields, including the economy. This development shows that the Internet is very supportive of activities in society. Starting from getting information to buying and selling goods online. One of the uses of the Internet in the economic sector is online trading. In this case, using the Internet is not only the relationship between buyers and sellers in buying and selling transactions but also includes activities to advertise and produce products or services.

* *Corresponding author, email address: fkartiasih@stis.ac.id*

The rapid development of internet networks, including in Indonesia, has encouraged the development of e-commerce in this country. The transition of conventional trading activities to online trading provides convenience for actors because there is no longer a time and place limit in its implementation. With the creation of this convenience, business actors have considerable business development opportunities to increase revenue. Apart from business actors, consumers, as one of the actors in trading or buying and selling activities, are also affected by the convenience caused, one of which is an increase in consumptive nature. The ease of transactions can cause this. The availability of goods is quite complete considering that there are no restrictions on places and locations that can limit the availability of goods and services, and also competitive prices with attractive promos are provided.

The COVID-19 pandemic has had an impact on social and economic activities. Economic activities that initially required direct interaction by sellers and buyers have turned into digital activities. Transforming economic activities to digital is the best alternative during this pandemic. According to data sourced from Bank Indonesia, the value of electronic money transactions increased by 8.63% from May 2020 to June 2020, indicating an increase in digital economic activities in Indonesia. The development of the digital economy is followed by innovations that support e-commerce activities to be faster, easier to use, and guaranteed security.

Every aspect of humanity and society is being transformed by the growth of e-commerce and the digital economy. It fundamentally changes how people communicate and interact in many domains, including public administration, business, education, and economic, political, and cultural growth. E-commerce thus offers opportunities and difficulties for everyone, especially those living in rural and remote places. E-commerce can be an opportunity, namely as a means to facilitate rapid, inclusive, and sustainable economic growth, improve living standards, and reduce poverty (Haji, 2021).

Statistics support the significant development of e-commerce in the global retail industry. According to estimates from the United Nations Conference on Trade and Development (UNCTAD), the global value of e-commerce reached 29 trillion USD in 2017. The Internet and other information and communication systems have room to contribute even more to this growing trend. The contribution of e-commerce to worldwide retail sales in 2018 was 12.2%. E-commerce accounted for 14.1% of all merchandise transactions worldwide in 2019.

The emergence of e-commerce opportunities encourages Micro, Small, and Medium Enterprises (MSMEs), which are a supporting factor for the economy in Indonesia, to be digitized with government support. One way is by conducting counseling or collaboration. MSME players need the transition to digital to survive in the current economic field. Research has proven that the number of Internet users and online sales by SMEs are indicators that provide the highest return to GDP (Fernández-Portillo et al., 2020). Starting from payment, bookkeeping, and sales, the MSME business system will be digitalized to achieve the target of 50 million Indonesian MSMEs entering the digital ecosystem. With the digitalization of MSMEs, the data generated from the digital process can be reused for analysis to develop the MSMEs themselves.

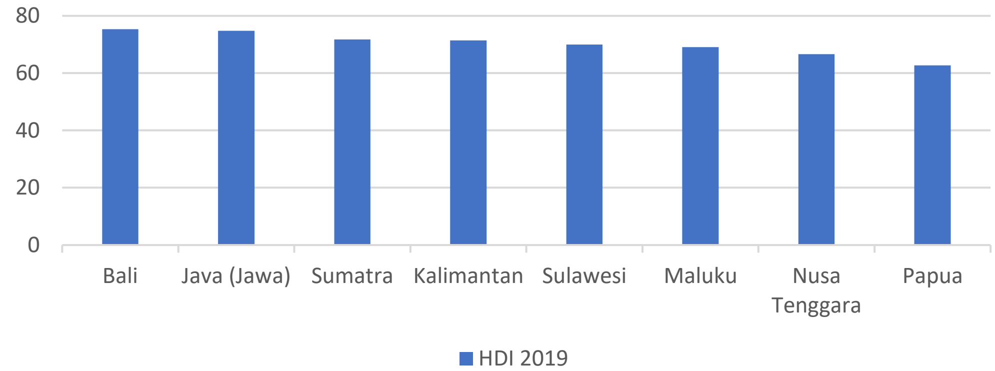
Sustainably and holistically structured e-commerce services can reduce environmental impact and often save costs (Van Loon et al., 2015). Companies can create service strategies to meet customers' actual needs while considering environmentally friendly options such as combining shipping and delivery of more maximized loads, using less impactful modes of transportation, and increasing pick-up points that also reduce packaging while lowering costs associated with service logistics (Muñoz-Villamizar et al., 2021). However, such solutions must be attractive to consumers.

Nowadays, all the needs of people in almost all fields can be obtained through e-commerce, including agriculture, mining, transportation, and others. In agriculture, ordering products can be done more easily, and the products available are more diverse. An example of e-commerce in agriculture is Tanihub. For the mining sector, there is Vmining, which helps sell coal. E-commerce, currently popular, includes buying and selling goods and services, restaurant services, and hospitality. E-commerce is also developing in the online transportation business; everything from ordering motorcycle taxis to ordering flight tickets can be done online. With the shift in buying and selling activities, the payment system has also been replaced from cash to credit, voucher, Gopay, and more. Finally, e-commerce also connects doctors with patients more quickly and efficiently in the health sector, such as Halodoc offers.

Economic growth can be seen based on GDP (Adwendi & Kartiasih, 2016; Kartiasih, 2019a; Kartiasih & Setiawan, 2020) through consumption and transaction activities in society, both e-commerce and conventional transactions. Several studies have been conducted and concluded that there is a positive relationship between e-commerce and GDP in Indonesia. Further, Dianary (2018) found that the effect of e-commerce development can be seen more clearly in the long term. Although Indonesia's Central Bureau of Statistics (BPS) has provided data related to e-commerce characteristics, few studies still use the data for further analysis.

Along with digital development, the availability of digital data is increasingly abundant, resulting in big data. One type of big data that has the potential to be utilized is Google Trends. Google Trends, a database provider platform containing the search frequency of a query on Google search services, can be used to see people's interest in something, including e-commerce. This is highly possible considering the tendency to use Google search service as the main search media nowadays. Seeing this potential, this research will also use Google Trends data as one of the data on e-commerce characteristics of a region.

The Human Development Index (HDI) significantly affects GDP (Arifin & Fadllan, 2021). Based on the HDI score available from BPS, which is shown in Figure 1, it is found that the Bali, Java, and Sumatra Islands have an average HDI value greater than other islands. A high HDI value shows the ability to understand something new well, one of which is technology. In 2019, based on Figure 2, Java Island dominated the percentage of e-commerce businesses higher than the other two islands, so it is interesting to know whether



Source: BPS (processed)
Figure 1. HDI value by islands in Indonesia in 2019

Source: BPS 2019 E-Commerce Statistics Publication
Figure 2. Choropleth map of the percentage of e-commerce businesses in 2019 in Java, Sumatra, and Bali Islands

the provinces on the island also have a more advanced economy and what e-commerce characteristics dominate in the groupings formed later. Given this, the research focuses on provinces on the three islands and the year 2019 for further analysis.

Based on some of the descriptions above and because no research discusses the characteristics of e-commerce that characterize GRDP, this study aims to determine the presence and absence of big data, to what extent e-commerce has on GRDP, and the grouping of its characteristics at the selected locus. This research focuses on the locus of 17 provinces on Java, Sumatra, and Bali islands. It utilizes BPS data as a source of Official Statistics and in the form of Google Trends as support.

Knowing the influence of e-commerce can help provide SME entrepreneurs a basis for developing e-commerce and a basis for government e-commerce policy. The characteristics of e-commerce based on GRDP groupings help development agencies, governments, consultants, and operating businesses better understand the contribution that e-commerce can make and improve the quality of e-commerce itself. This helps them to prioritize, plan, and evaluate e-commerce projects.

2. LITERATURE REVIEW

Economic sector analysis is important, especially in the context of community welfare. Improved community welfare can be reflected through economic growth, measured using GRDP indicators. In the era of technological development, the use of the Internet is increasingly penetrating the economic sector, with E-commerce being one of its manifestations. The rapidly growing e-commerce has the potential to impact the economy of a region or country significantly. In addition, the widespread adoption of the Internet also enriches digital data, such as the relative frequency of certain searches on search engines, as can be found on Google Trends. The characteristics of e-commerce, including the patterns of search words or queries reflected in Google Trends, offer further insights. The diversity of E-commerce characteristics and the variation in the level of economic growth in each region provide a basis for clustering aimed at identifying distribution patterns and relationships between attributes.

2.1. Gross Regional Domestic Product (GRDP)

The economic condition of a region is one of the most important things that will support the welfare of the region's residents. Economic conditions are also expected to develop along with the development of community needs. At the regional level, this is a change in the economic conditions of a region for the better. The increase in community prosperity caused by the development of economic activities in a region can be represented by economic growth (Hawari & Kartiasih, 2017; Kusumasari & Kartiasih, 2017).

Regional economic growth can be measured by components or indicators representing the state of the regional economy, such as GRDP. According to BPS, the total value added generated by all business units in a region is reflected in the Gross Regional Domestic Product (GRDP). GRDP, then, represents the sum of final goods and services produced in all sectors of the regional economy. This notion provides a broad perspective on economic activities and how many industries contribute to the region's economic growth. GRDP can be used to assess the state of the economy at both current and fixed prices. The value added to products and services is estimated in the GRDP based on current prices and thus can be used in assessing the economy's structure in a particular year. While GRDP, which is based on constant prices, assesses the value added using prices from a particular year, it can also be used to track the region's economic growth over time without being affected by price changes occurring in that year.

Based on the previous explanation, it can be concluded that GRDP is an important indicator used to assess the state of the economy of a region. In addition, GRDP can also be used to compare regional economic performance (Kartiasih, 2019b; Ningsih & Kartiasih, 2019; Subian et al., 2024). Governments, businesses, and policymakers can support sustainable economic growth and improve living standards in the region by taking strategic actions based on their understanding of GRDP.

2.2. E-commerce

The Internet has opened up new opportunities for people to develop and live a more practical life. Electronic commerce, or e-commerce, is one of the observable effects of today's technological development. E-commerce greatly benefits consumers and sellers by allowing buying and selling activities to be carried out online without limiting place and time.

E-commerce is any trading activity or business interaction involving Internet technology (Mahir, 2015). In addition, e-commerce is also defined by trade, which refers to the extensive use of digital technology in facilitating online business processes such as sales and transactions (Mashalah et al., 2022). Marketplaces such as Lazada and Shopee are examples of e-commerce business models that serve as platforms to connect sellers and buyers. Sellers only need to focus on buying and selling activities and fulfilling orders, while site management is the responsibility of the platform owner. There are four types of activities in e-commerce.

1. B2B (Business to Business)
   B2B is the activity of buying and selling goods or services in e-commerce between companies. In this business, transactions generally involve large amounts of funds and products, for example, Electronic City.
2. B2C (Business to Consumer)
   B2C is the activity of buying and selling goods or services on e-commerce from companies to potential customers. The difference with B2B, in this business, business people usually sell products at retail, for example, e-commerce that utilizes the Lazada, Blibli, and Shopee platforms.
3. C2C (Consumer to Consumer)
   C2C is the activity of buying and selling goods or services in e-commerce between individual consumers. An example that conducts this type of activity is e-commerce, which utilizes the OLX and Kaskus platforms to resell used goods to other consumers.
4. C2B (Consumer to Business)
   C2B is the activity of buying and selling goods or services on e-commerce from individuals to companies. An example of e-commerce in this type of activity is a freelancer, a person with special expertise or skills, such as a graphic designer, who offers his services to create a logo for a client company or a writer who creates writing for publishers.

2.3. Clustering

Clustering is a process in which similar objects are grouped. By applying clustering techniques, we can recognize the density and distance between regions in object space and comprehensively discover distribution patterns and relationships between attributes. The classification approach can also be an effective tool to distinguish groups or classes of objects. Clustering is a data mining technique that is carried out to group large data into groups with similar characteristics in each group. It can be said that clustering is an unsupervised learning method because the basis of the grouping uses the similarity of the objects studied instead of learning from similar data labels that already exist (Nugraha et al., 2014).

Hierarchical and non-hierarchical clustering are the two main types of clustering. A data clustering technique called hierarchical clustering starts by combining two or more objects that are most comparable to each other. Other objects with the second closest degree of proximity will be added to this process, and so on. As a result, the cluster will create a tree with different hierarchies, with the most similar things on one branch and the most different objects on the other branches. Eventually, all objects should be connected into a single cluster. Dendrograms are a common visual aid to illustrate this hierarchical process. On the other hand, non-hierarchical clustering approaches, in contrast to hierarchical methods, start by first determining how many clusters are desired, such as two, three, or more clusters. Once the number of clusters is set, the clustering process proceeds without adhering to a hierarchical structure.

A non-hierarchical clustering technique called K-Means seeks to divide data into several clusters containing data with similar properties (Benri et al., 2015). This strategy assigns data to the closest cluster center using a distance-based methodology. K-Means is renowned for its effectiveness in classifying very large volumes of data and finding outliers. Using K-Means, data is partitioned into interrelated groups based on similar characteristics (Kartiasih et al., 2023, 2023a, 2023b). This method does not involve a hierarchy between clusters but instead focuses on forming clusters independently. It allows for simpler structure identification and the ability to pre-determine the desired number of clusters.

2.4. Google Trends

As analog information media transforms into digital, people live in the digital era (Yulianti et al., 2021). In addition, an increasing amount of social data has been captured digitally. The term that emerges in this context is big data. These data also illustrate the concept of big data, which is a very large volume, a large variety of data at high speed, and complex diversity (Pujianto et al., 2018). The abundance of data today offers much promise for use in various contexts. One type of big data that can be utilized is Google Trends.

Google Trends is a tool from Google Inc. used to determine searches and search patterns on the Google platform for a certain period (Purnaningrum & Ariqoh, 2019). Google Trends collects data from various sources, including web searches, images, YouTube, news, and Google Shopping, to provide insights into trending trends. In addition, Google Trends also provides a wide array of categories covering various topics such as Arts & Entertainment, Tourism, Sports, Work & Education, Scientific, Health, Finance, and many more. Thus, users can explore the latest trends and information in various fields.

Google Trends is a database provider platform that displays the relative frequency of a search term (query) to the total search volume in various regions worldwide (Totleben & Kuźmar, 2022). In other words, Google Trends allows users to see how often a keyword or topic is searched by internet users from various geographical locations. This data provides an overview of the popularity and growing trends related to a particular topic in the global community. By accessing Google Trends, users can gain valuable insights into search interests and trends in different regions and identify trend patterns that are useful for research, marketing, or broader market understanding purposes. Using Google Trends as a data source or other utilization is highly potential. This is due to the tendency to use Google search services as the first choice when searching the Internet (Purnaningrum & Ariqoh, 2019). Therefore, in this research, Google Trends will be used as one of the data sources.

Based on a systematic literature review, this study developed the use of clustering to see the characteristics of e-commerce in GRDP groupings. Several previous studies, such as Aula & Suharto (2021), Dianary (2018), Karina et al. (2021), Nasution et al. (2020), Pantelimon et al. (2020), Khotimah (2021) and Siregar (2019), which analyze the economic sector, especially related to economic growth, provides a view of the importance of the economic sector so that it needs to be studied and analyzed. These studies provide a useful perspective in this research because they provide an overview of the relationship between e-commerce and the economy. Khotimah's (2021) and Siregar's (2019) research suggested a relevant view because they also focus on analyzing the economic sector, especially economic growth, with cluster analysis. They used a similar method in clustering, namely K-Means, to conduct cluster analysis in the economic sector to provide references and guidance in approaches that have proven effective in developing clustering methodologies in this study. The research results of Giswandhani & Hilmi (2020) and Jamaludin (2022) are the basis for determining several e-commerce characteristic variables tested in this study.

In contrast to previous research studies, this study specifically focuses on using clustering methods to analyze e-commerce characteristics and relate them to Gross Regional Domestic Expenditure (GRDP). This makes an important contribution to understanding the characteristics of regions with different GRDPs. This research also shows uniqueness by integrating previous research, using Google Trends data, and contributing to developing clustering methodology in economic sector analysis and e-commerce. By identifying e-commerce characteristics related to economic growth, this study provides valuable information and deeper insights into the relationship between GRDP, economic sectors, e-commerce, and economic growth at the provincial level.

3. METHODOLOGY

3.1. Data

According to BPS, data is the result of observation, measurement, or collection of information from various sources, such as facts, figures, or other types of information. Data can be represented numerically, textually, graphically, or sonically and processed into more valuable and relevant information. To ensure the veracity and dependability of data, a methodical and measured approach is usually used during the data collection process.

In this study, the data used is quantitative and sourced from secondary data or data that is not collected independently but from the collection of certain institutions and has been published. Secondary data is often used in research because it can provide extensive information and cover a large sample size without spending time and money collecting new data. However, researchers must ensure the secondary data is correct, valid, and relevant to the work or analysis.

Official statistical data from BPS and big data are the secondary data used in this investigation. These include the GRDP Growth Rate per province at constant 2010 prices for 2019 from website BPS and e-commerce characteristics data for 2019, taken from the BPS publication "Statistik E-commerce 2019". To obtain true results (in the sense of ceteris paribus) and isolate the effects of the variables under investigation, 2019, one year before the pandemic started, was chosen as the relevant period. The publication "Statistik E-Commerce 2019" is a document prepared based on data collection results from the 2019 E-Commerce Survey, which was carried out by listing a sample of selected census blocks in all provinces in Indonesia (BPS, 2019). This publication survey uses a systematic sampling method using the SE2016-L2 sample frame. In this research, the published data will focus on 17 provinces on three islands: Java, Sumatra, and Bali.

The big data used is Google Trends data. Google Trends is a statistical and graphical tool that can display trends in search topics on the Google search engine over a customizable period. Google Trends presents information through visualizations that allow users to understand what keywords or topics are most searched at the moment. The tool allows users to see search trends in real-time and compare specific keywords or topics in different periods. Google Trends also compares each province's interest in a keyword. In addition, this technology allows users to forecast future search trends and base judgments on facts and knowledge obtained from various sources.

This research uses Google Trends to find data on public interest in Indonesia for a keyword in each province in 2019. The keywords used are words related to e-commerce. The transactions or buying and selling of products or services through electronic media or online can be called e-commerce (Maulana et al., 2021). The process involves consumers or customers buying products from companies or sellers and business transactions between companies carried out using computers as intermediaries. In this case, the purchase or sale of products is done online through a website or e-commerce platform provided by the company or seller. The first keyword used is "e-commerce" itself. From the explanation of e-commerce above, keywords that are closely related or synonymous with the word e-commerce are selected. In addition, the keywords chosen have few syllables, so the scope of data detected by Google Trends is wider so that it is more representative. From this explanation, the next keywords used are "sell online" and "buy online". E-commerce is closely related to applications that facilitate online buying and selling interactions. Based on Putri & Zakaria (2020), the top 3 e-commerce platforms in Indonesia are Shopee, Tokopedia, and Bukalapak. Therefore, the keywords used in Google Trends search are "e-commerce", "jual online" (Eng: sell online), "beli online" (Eng: buy online), "Shopee", "Tokopedia", and "Bukalapak". The list of variables used in this study is presented in Table 1.

Table 1. List of variables

Variables	Concepts and Definitions	Source
Growth Rate of Gross Regional Domestic Product at Constant 2010 Prices	Gross Regional Domestic Product (GRDP) measures the difference between the total value of goods and services produced in a region in the current and previous periods. The Central Bureau of Statistics (BPS) states that GRDP is the sum of added value generated by all business entities or economic sectors, both production factors originating from within and outside a particular region. GRDP at constant prices, which is determined by using prices from a particular year as the base year, shows the added value of these goods and services without being affected by inflation/changes in prices. This variable was chosen because it covers the purpose of seeing economic development and has the same units as the e-commerce characteristics to be studied, namely percent.	BPS
Percentage of E-commerce Businesses by Province	The percentage of establishments within a province that accept or sell products or services online.	BPS

Variables	Concepts and Definitions	Source
Percentage of E-commerce Businesses by Province and Business Field. (Business fields covered: Agriculture, Forestry, and Fisheries (A); Mining and Quarrying (B); Manufacturing (C); Construction (F); Wholesale and Retail Trade, Repair and Maintenance of Automobiles and Motorcycles (G); Information and Communication (J); Financial and Insurance Activities (K)	In the available data, there are several categories of business fields, but in this study, these categories will be grouped into two based on the specified criteria. Based on BPS data, the median of the average GRDP of business fields in each province in 2019 was 22,166.12 billion rupiahs. Therefore, two groups were formed: the category of business fields above and below the median average contributor to GRDP. In this study, the group used is the group with business fields contributing to GRDP above / equal to the median value and obtained seven business fields mentioned in the variable column. This is done to determine whether the business field category that contributes the highest GDRP conventionally is also included in the high GDRP group when run through e-commerce.	BPS
Percentage of E-commerce Businesses by Province and Employment (variables with employment of 1-4)	This data is intended to see the category of GDP growth based on the smallest number of e-commerce workers (BPS). This is due to the characteristics of e-commerce sales that do not require so many workers because they are engaged in the online field.	BPS
Percentage of E-commerce Businesses by Province and Total Revenue Value (Total Revenue Value Covered: 300 Million Rp-2.5 Billion Rp; 2.5–50 Billion Rp; >50 Billion Rp)	The highest percentage proportion of e-commerce businesses by total revenue is in the <300 million category. However, several studies have concluded that total revenue positively affects GDP. Therefore, the e-commerce revenue data used is the above 300 million categories to see the characteristics of GDP by income category covered by a smaller proportion.	BPS
Percentage of E-commerce Businesses by Province and Frequently Used Payment Method: (Payment methods covered are Cash/COD)	Based on Giswandhani & Hilmi (2020), non-cash payments positively affect the consumptive nature of society. This is also related to GDP in terms of household expenditure. However, in BPS data in 2019, the percentage of e-commerce payment methods is still dominated by cash or COD payments. Therefore, this study chose the COD payment variable to see whether e-commerce with this method falls into the high or low GDP category.	BPS
Percentage of E-commerce Businesses by Province and Transaction Value	The highest percentage proportion of e-commerce businesses by transaction is in the <300 million category. However, several studies have concluded that transaction value positively affects GDP. Therefore, the e-commerce transaction value used is above 300 million to see the characteristics of GDP based on the transaction value category covered in a smaller proportion.	BPS
E-commerce interest by province. There are 6 variables with each of the following keywords "e-commerce"; "beli online"; "jual online"; "shopee"; "tokopedia"; "bukalapak"	To see the characteristics of GRDP based on public interest, search keywords on Google Trends related to e-commerce were used in this study.	Google Trends

Source: Statistik E-Commerce 2019 (BPS Publication) & Google Trends

In Google Trends, the data scale used to describe the popularity or interest of a topic is relative rather than an absolute number. The data is measured on a scale of 0 to 100 and represents the relative level of interest in a topic or search term over a period of time. The 0-100 scale in Google Trends does not indicate an absolute number or percentage but rather indicates the extent to which a topic or search term is relatively popular within a specified period.

3.2. Methods

Figure 3 shows the flow of methods used to analyze the data that has been collected to meet the research objectives.

3.2.1. Regression Analysis

One dependent variable (Y) and one or more independent variables (X) are tested for their relationship by forming a linear regression mathematical model. The direction and degree of influence of the independent factors on the dependent variable are investigated using this method. In general, the relationship can be expressed as the following equation:

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon \qquad (1)$$

Where X is the independent variable, Y is the dependent variable, β_0 and β_1 are known as regression coefficients in the form of parameters whose values are unknown, and ε is a random error. This study will use the percentage of e-commerce businesses and GRDP in a particular region in the linear regression analysis. The regression coefficients (β_0 and β_1) will be used to measure the effect of the independent variable (percentage of e-commerce businesses) on the dependent variable (GRDP), while the random error (ε) will illustrate how accurate the regression model is.

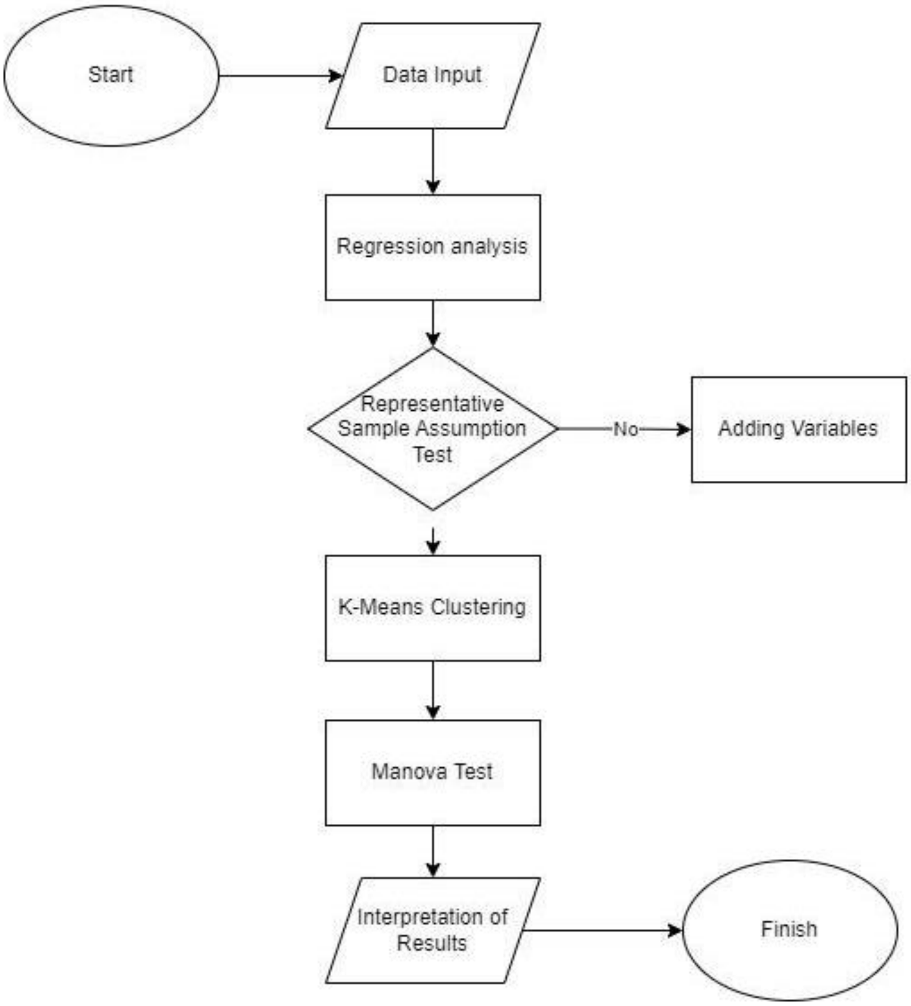


Figure 3. The flow of research method

3.2.2. Cluster Analysis

Cluster analysis is a statistical method for grouping research objects according to their properties or characteristics (Awalluddin & Taufik, 2017). The main purpose of cluster analysis is to form groups or clusters with similar characteristics internally but with differences or diversity of characteristics externally. In cluster analysis, the variables used to assess the similarity between objects must be well considered to produce an optimal solution. In addition, sample representation also needs to be considered so that the cluster analysis results are reliable. Cluster analysis is not a statistical inference technique that estimates population parameters based on a sample. Instead, cluster analysis measures the structure of a group of observations. Therefore, it has a strong mathematical basis but not a statistical basis. Important criteria such as normality, linearity, and homoscedasticity in other techniques have little effect on cluster analysis. However, researchers need to pay attention to sample representation issues.

The assumptions of representative sample cluster analysis are as follows. Most researchers cannot access all the population data required for cluster analysis. Therefore, they must rely on a representative sample of cases from the larger population and create clusters based on that sample. For cluster analysis, population data is rarely available to researchers. Therefore, a sample of cases is used instead of population data. However, researchers must ensure that the sample represents the population structure. The outliers in the sample may belong to different groups, so removing them will result in a bias in the structure estimation. To verify that the cluster analysis results can be applied to the target population, researchers must ensure that their sample is representative. Although cluster analysis is useful in grouping research objects, the results are only as good as the representativeness of the sample used.

The Kaiser-Meyer-Olkin (KMO) test can be used to conduct a test to ensure the sample taken is representative of the population of data being analyzed. KMO evaluates the overall feasibility of the sample as well as the practicality of each indicator in this test. A scale from 0 to 1 is used to express the findings of the KMO test conducted. The sample collected is considered representative of the population if the KMO value exceeds 0.5.

Several techniques can be selected and performed to perform data clustering. K-Means is one of the strategies used to categorize data with a partitioning system and is an unsupervised data analysis or data mining method. With this method, data must go through a normalization or standardization stage if it has different units or a range that is too far between variables. In cluster analysis, striking differences in units can lead to inaccurate results. Standardization can be done with a Z-score. This standardization value is obtained by subtracting the data from its mean and dividing by its standard deviation. Furthermore, the data is processed with the K-means algorithm as described below.

1. Determine the optimal k or how many clusters to form.
2. Grouping random data into clusters.
3. Using the existing data in each cluster, determine the cluster center (centroid) using the formula:

$$C_{kj} = \frac{X_{1j}+X_{2j}+\cdots+X_{nj}}{n} \quad (2)$$

   Where C_{kj} is the kth cluster center on the jth variable (j=1, 2, ..., p), and n is the amount of data in the kth cluster.

4. Calculate the distance between each object and each centroid using Euclidean distance to find the distance between each object and each centroid.

$$d(x_i, x_g) = \sqrt{\sum_{j=1}^{p}(x_{ij} - x_{gj})^2} \quad (3)$$

5. Use the formula to calculate the objective function:

$$J = \sum_{i=1}^{n}\sum_{j=1}^{k} a_{ij} d(x_i, C_{kj})^2 \quad (4)$$

6. Each data item is placed to the nearest center of mass or average, which is expressed as follows:

$$a_{ij} = \begin{Bmatrix} 1, & s=\min\{d(x_i, C_{kj})\} \\ 0, & others \end{Bmatrix} \quad (5)$$

7. Until the object does not move or the objective function does not change, repeat steps 3-6.

3.2.2.1. Determination of the number of clusters

By comparing the result of the number of clusters to be generated with the sum of squares error (SSE) value, the Elbow method is one way to determine the ideal number of clusters. This technique can be determined through a graph that shows the angle at a certain point. The angle is the result of comparing the value of the first cluster and the value of the second cluster that has decreased the most. The SSE value of each cluster can be calculated to get a comparison. The SSE value decreases as the number of clusters increases.

$$SSE = \sum_{K=1}^{K} \sum_{i=1}^{n} |x_i - C_k|^2 \quad (6)$$

Description:

K = k-th cluster
X_i = distance of data to the i-th object
c_k = i-th cluster center

The Silhouette Coefficient is an evaluation technique to assess the quality and strength of clusters in a cluster analysis by measuring how good or bad an object is in a particular group. This method combines the concepts of separation and cohesion in cluster evaluation. For each data object, the Silhouette index value must be calculated to determine the Silhouette coefficient value. The largest Global Silhouette Index value is between 2 and n-1 clusters, where n is the total number of data objects in the cluster analysis and is used to generate the Silhouette coefficient value.

$$SC = max_k \; SI(k) \quad (7)$$

Description:

SC = Silhouette Coefficient
SI = Silhouette Global Index
k = number of clusters

The Dunn validity index (DN) is a cluster quality evaluation method in cluster analysis. It compares the maximum value of cluster diameter as compactness and the value of dissimilarity function between the two clusters as separation. The quality of the resulting cluster increases as the DN value increases. The difference between the dissimilarity function value of the two clusters and the highest cluster diameter value is used to determine the DN value. Thus, the larger the DN value, the better the number of clusters can be determined.

3.2.2.2. Manova

The Manova test was applied to ensure that the resulting clusters were significantly different from each other. A statistical method called multivariate analysis of variance (Manova) is used to determine whether significant differences exist between the means of many variables from many populations at once. Manova measures the difference between two or more metric-dependent variables based on several non-metric independent factors. Manova is used to test the relationship between the dependent variable and the independent variables in a study and can be used to measure how similar the mean vectors of multiple populations are to each other.

Hypothesize as follows:

$H_0: \mu_1 = \mu_2 = \cdots = \mu_i$
$H_1: There\ is\ at\ least\ one\ \mu_i\ \ that\ is\ not\ the\ same\ as, i = 1,2,3, \ldots, g$

The test statistic used is Wilk's Lambda Λ∗, which can be calculated using the following formula:

$$\Lambda^* = \frac{|W|}{|B+W|} \quad (8)$$

$$W = \sum_{j=1}^{g} \sum_{i=1}^{nl} (x_{ji} - \bar{x}_j)(x_{ji} - \bar{x}_j)' \quad (9)$$

$$B = \sum_{j=1}^{g} n_j (\bar{x}_j - \bar{x})(\bar{x}_j - \bar{x})' \quad (10)$$

B and W are the sum of squares and cross-product matrices between and within groups, respectively, with

degrees of freedom g-1 and Σni-g.
Description:

X_{ji}: vector of i-th observation in group j

X_j: jth group mean vector

n_j: number of group individuals in the jth group

$\bar{x}$: average observation vector of all groups

Wilk's Lambda value can be found with the F test statistic. The criterion is to reject H0 if lambda > F table or p-value < alpha, which means there is a difference in means between groups/clusters.

4. RESULTS AND DISCUSSION

In 2019, the percentage growth rate of GRDP in Java, Sumatra, and Bali showed that the Special Region of Yogyakarta province has the highest value. On the other hand, Riau province has the lowest GDP rate value. In fact, it was also found that the percentage of e-commerce in the two provinces follows the same pattern, with DIY having the highest percentage and Riau having the lowest percentage. This suggests that descriptively, the percentage of e-commerce in a province may affect GRDP. To verify the hypothesis, it is necessary to perform a regression test on the data. The results can be seen in Table 2.

In the regression analysis, two hypotheses are proposed: the null hypothesis is that e-commerce businesses do not affect GRDP, and the alternative hypothesis is that e-commerce businesses affect GRDP. Based on the test results, it was found that the p-value is less than alpha (20%) or rejects H0. This means that the percentage of e-commerce businesses positively affects GRDP in Java, Sumatra, and Bali, with an 80% confidence level. This result in Table 2 expands the scope of Khotimah (2021), which stated that there is a positive influence of e-commerce on the economic growth of provinces in Java.

The regression equation model is as follows.

$$\widehat{PDRB} = 2.9391 + 0.1288\, ecommerce$$

The regression model above can be interpreted that with every one percent increase in e-commerce percentage, the GRDP growth rate increases by 0.1288%. In addition, the R-squared value of 0.4402 indicates that 44% of the 2019 GRDP growth rate can be explained by the percentage of e-commerce businesses in 2019.

Residual Normality Assumption Testing

The regression model that has been formed is then tested for residual normality assumptions. This model has a null hypothesis, namely normally distributed residual data, and hypothesis one, namely residual data, is not normally distributed. From the test results, as can be seen in Table 3, it is found that the statistical test results for both Shapiro Wilk, Kolmogorov Smirnov, and Jarque Bera have a p-value of more than 20% alpha, so they failed to reject H_0. It indicates that, with 80% confidence, the regression model residuals are normally distributed.

Table 2. Regression analysis results

Variables	Coefficient	Std. Error	t-value	p-value
Intercept	2.9391	0.6201	4.740	0.000263
E-commerce	0.1288	0.0375	3.434	0.003690
R^2 = 0.4402				

Table 3. Results of residual normality assumption test of regression model

Equation	Shapiro-Wilk		Kolmogorov-Smirnov		Jarque-Bera	
	Statistics	p-value	Statistics	p-value	Statistics	p-value
Linear Regression	0.97316	0.8711	0.11284	0.9648	0.18157	0.9132

Kaiser-Meyer-Olkin (KMO)

After identifying the positive relationship between e-commerce and GRDP, the study seeks to determine what characteristics characterize the grouping of GRDP values in Java, Sumatra, and Bali provinces. The characteristics of e-commerce used in characterizing GRDP include business field, number of workers, revenue value, transaction value, and payment methods provided. In addition, big data through Google Trends data is also included in the grouping of GRDP characteristics. This is used to see the characteristics of GRDP based on the search for a keyword related to e-commerce. This data can be used as a representation of public interest in e-commerce based on the size of the keyword search. In the initial stage, data standardization was carried out. This is because there are differences in units between the variables used. Furthermore, Kaiser-Meyer-Olkin (KMO) testing is carried out to test the adequacy of the sample used. The following are the results of the KMO test.

The test results show that some variables have an MSA KMO value of less than 0.5 (see Table 4). These variables include the Selected Business Field, COD Payment, Keyword "Jual Online", Keyword "Beli Online", and Keyword "Tokopedia". However, these four variables are retained and used in this study. This is because the overall MSA KMO is worth 0.51 or more than 0.5 so that the sample of e-commerce characteristics based on selected BPS data and big data together can be said to represent the population or be representative enough, and the analysis can be continued.

K-Means Cluster Analysis

To see what e-commerce characteristics reflect the GRDP, a K-Means cluster analysis was conducted to form a group containing GRDP and these characteristics. The first step is determining the optimal number of clusters (K). The following are the results of several tests to determine the optimal K.

Table 4. KMO results

Overall MSA = 0.51					
MSA of Each Variable =					
Selected Business Fields	Number of Workers (1-4 People)	Revenue (>300 million)	Transaction (>300 million)	COD Payment	Keywords "e-commerce"
0.48	0.66	0.57	0.55	0.44	0.62
Keywords "Jual Online"	Keywords "Beli Online"	Keywords "Shopee"	Keywords "Tokopedia"	Keywords "Bukalapak"	
0.20	0.38	0.64	0.49	0.61	

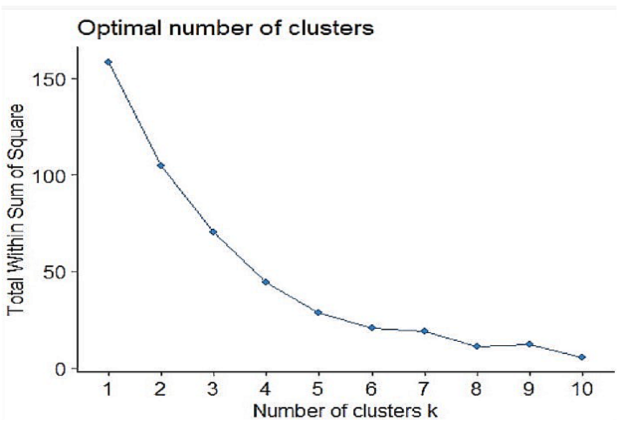


Figure 4. K Optimal Elbow method

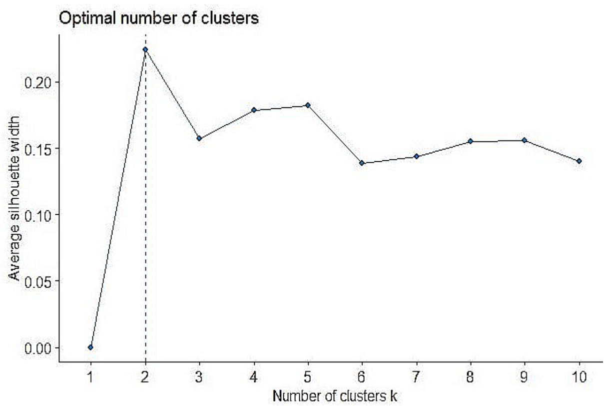


Figure 5. K Optimal Silhouette method

Table 5. Optimal K testing results

	Value	Methods	Cluster
Connectivity	2.9290	k means	2
Dunn	0.8992	k means	2
Silhouette	0.4443	k means	2

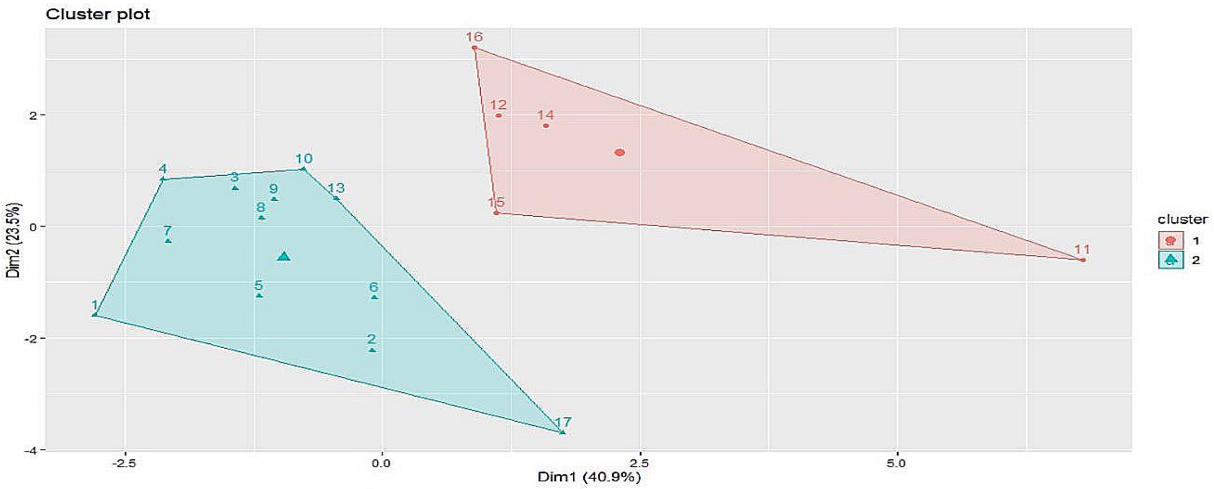


Figure 6. Cluster results

Based on the graphs from the Elbow (see Figure 4) and Silhouette methods (see Figure 5), as well as the Dunn, Connectivity, and Silhouette tests, the majority of the optimal K for dividing GRDP in Java, Sumatra, and Bali, is 2 clusters (see Table 5). So, based on these results, two cluster models were formed. The results of the clustering of the two clusters can be seen in Figure 6. From the figure, the groups formed have shown good clustering results because the two groups are separated from each other. This shows that the variables within one group have homogeneous variations, while the variables between groups show heterogeneity.

Profiling Cluster Results

Cluster 1

Based on Table 6, cluster 1 shows a higher average GRDP than cluster 2. Therefore, cluster 1 is called the high GRDP cluster. In this cluster, the e-commerce characteristics of provinces on the islands of Sumatra, Java, and Bali that tend to be higher than cluster 2 include a total income value above 300 million rupiahs and a transaction value above 300 million rupiahs. This is under the existing conditions, namely that if transactions and income are high, then the GDP in a region will also increase. On the other hand, e-commerce characteristics that tend to be lower in this cluster include the following.

1. Business field
   The selected business sectors are Agriculture, Forestry and Fisheries (A); Mining and Quarrying (B); Manufacturing Industry (C); Construction (F); Wholesale and Retail Trade, Repair and Maintenance of Cars and Motorcycles (G); Information and Communication (J); Financial and Insurance Activities (K). These businesses are characterized by high GRDP when run offline. However, from the clustering results, these businesses, when conducted through e-commerce, are no longer characterized by high GRDP. This may occur because these businesses tend to be conducted conventionally.
2. Labor force of 1-4 people
   Although e-commerce is online-based and generally does not require a large number of workers, the cluster results show that clusters with high GRDP are still characterized by the number of workers above five people. It can be said that there is still no change in the trend of the contribution of the number of workers to GRDP in both e-commerce and conventional businesses. That is, a large number of workers is also characterized by high GRDP.
3. Payment method in person (COD)
   Currently, there are indirect payment services such as mobile money (Ovo, Dana, LinkAja), online vouchers, online credit and credit cards (Kredivo, Akulaku), debit cards or online bank transfers,

online payment services (PayPal, GooglePay), points from reward programs, and so on. From the clustering results, these indirect payments tend to have a large contribution in characterizing high GRDP in an area compared to the results of direct payments (COD). This is in line with Giswandhani & Hilmi (2020), who argue that the dimension of ease in using non-cash transaction tools significantly influences people's consumptive behavior.

In addition, based on big data, namely searches through Google Trends, all keywords chosen to describe GDP in 2019 for each province on the selected island show a tendency for higher searches compared to cluster 2. This can be attributed to the high public interest in e-commerce, which also shows a tendency for high GDP in the area. When analyzing the clustering results by region, the provinces included in the high GRDP cluster tend to cluster on the island of Java. These provinces include DKI Jakarta, West Java, DI Yogyakarta, East Java, and Banten. This can be seen from Figure 7. The dominance of Java Island can be a sign that there is still a gap in knowledge or dissemination of e-commerce utilization by the community in the three islands.

Cluster 2
Cluster 2 is called the lower GRDP cluster because it has a lower average value of GRDP growth rate than cluster 1. It is presented in Table 6. The e-commerce characteristics in this cluster that tend to be lower than cluster 1 are the total revenue and transaction value of less than 300 million. However, this cluster has higher e-commerce characteristics, such as a business field, a workforce of 1-4 people, and a direct payment method (COD).

Table 6. The average value of variables for each cluster

Variables	Cluster	
	1	2
GRDP	5.64%	4.71%
Selected Business Fields	55.33%	60.03%
Number of Workers (1-4 People)	81.76%	84.79%
Revenue (>300 million Rp)	19.74%	15.04%
Transaction (>300 million Rp)	12.07%	7.50%
COD Payment	77.96%	82.91%
Keyword "E-commerce"	81.4	61
Keyword "Jual Online"	59	54.08
Keyword "Beli Online"	68	57.67
Keyword "Shopee"	86.8	63.08
Keyword "Tokopedia"	71	32.83
Keyword "Bukalapak"	82.6	50.42

Note: The Google Trends keyword variable is an index with a range of 0-100

Figure 7. Choropleth map of e-commerce characteristics grouping results

Table 7. Manova results

	Df	Wilks	Approx F	Num DF	DF Den	p-value
kmRes.cluster	1	0.089151	3.4056	12	4	0.1235*
Residuals	15					

Note: *) significant at 20% alpha.

In addition, based on big data, namely Google Trends searches, all keywords used to describe GRDP in 2019 in each province on the selected island show a lower search trend than cluster 1. In this cluster, 12 provinces are included in the lower GRDP cluster, namely Aceh, North Sumatra, West Sumatra, Riau, Jambi, South Sumatra, Bengkulu, Lampung, Bangka Belitung Islands, Riau Islands, Central Java, and Bali.

Manova

Because there are 12 variables in each cluster, the Manova test was conducted to check that the two clusters' groupings differed. The hypotheses in this test include the null hypothesis, namely that there is no difference in the average characteristics between the two clusters, and hypothesis one, namely that there is a difference in the average between the two clusters. From the Manova test results in Table 7, it can be seen that the p-value <0.2 (alpha) so that the conclusion is obtained to reject H_0. Therefore, through the Manova test above, it can be concluded that with a confidence level of 80%, there are differences in characteristics between clusters on the response, which in this case is the characteristics of e-commerce and GRDP conditions in the provinces in Java, Sumatra, and Bali.

5. CONCLUSIONS AND SUGGESTIONS

Based on the research that has been conducted, the conclusion is that the percentage of e-commerce businesses has a positive effect on the GDP growth rate on the islands of Java, Sumatra, and Bali, with a contribution of 44%. This can be a basis for MSME owners to expand their operations into e-commerce and a means for the government to encourage the digitalization of MSMEs so that the number of e-commerce businesses in the region increases. In addition, clustering using the K-Means method produces two optimal clusters and fulfills the Manova test, indicating that the two clusters have been proven to be different. In the high GRDP cluster, the supporting e-commerce characteristics include high transaction value and revenue value and high keyword searches related to e-commerce in the area. The regions that fall into this cluster are dominated by provinces in Java, except Central Java. This condition shows that there is still a gap between Java Island, Sumatra Island, and Bali Island in terms of e-commerce utilization to support the increase of GRDP in each province. On the other hand, the lower GRDP cluster has high e-commerce characteristics, including business field, workforce of 1-4 people, and direct payment method (COD). The regions included in the lower GRDP cluster include 12 provinces, namely Aceh, North Sumatra, South Sumatra, West Sumatra, Riau, Jambi, Bengkulu, Lampung, Bangka Belitung Islands, Riau Islands, Central Java, and Bali.

The results of this grouping can provide knowledge regarding which provinces are included in the low GRDP group based on their e-commerce characteristics. These regions can make various evaluations and improvements in e-commerce to increase their GRDP. The first thing that can be done is to create plans and policies to increase the number of transactions and revenue originating from e-commerce and other policies that can increase GRDP through e-commerce. The design of this plan can be based on the grouping of e-commerce characteristics, which provides knowledge regarding business fields that have a high contribution to GRDP but apparently do not make it into the characteristics of high GRDP when implemented through e-commerce. This information can be used as material for consideration to maximize this field through e-commerce or change the focus of e-commerce development to other fields. The design can also be based on considering the number of workers, digitalizing payments, and expanding the dissemination of information about e-commerce, which is expected to increase public interest in e-commerce.

Furthermore, there are several ways that policymakers can improve e-commerce in their regions. Following the example of other successful countries, the Indian government has made various efforts, such as allocating funds for projects to facilitate high-speed internet access and wifi hotspots in villages and providing low tariffs for digital services (Anuj et al., 2018). In addition, the performance of existing e-commerce can also be improved to support GRDP, such as by encouraging people to make payments by indirect methods and increasing the sales capacity of e-commerce so that it can absorb more labor.

This research can still be carried out with various improvements, such as adding variables, choosing more Google Trends keywords so that they can better describe the actual conditions, and the use of other big data sources. This research has development opportunities regarding the coverage of the area studied. This opportunity can be by expanding the coverage area or focusing on deepening a smaller area in the research so that the policy or understanding obtained is broader and more detailed.

REFERENCES

Adwendi, S. J., & Kartiasih, F. (2016). Penggunaan Error Correction Mechanism dalam Analisis Pengaruh Investasi Langsung Luar Negeri Terhadap Pertumbuhan Ekonomi Indonesia. *Statistika: Journal of Theoretical Statistics and Its Applications*, *16*(1), 17–27. https://doi.org/10.29313/jstat.v16i1.1767

Anuj, K., Fayaz, F., & Kapoor, M. N. (2018). Impact of E-Commerce in Indian Economy. *IOSR Journal of Business and Management (IOSR-JBM)*, *20*(5), 59–71. https://doi.org/10.9790/487X-2005065971

Arifin, S. R., & Fadllan. (2021). Pengaruh Indeks Pembangunan Manusia (IPM) dan Tingkat Pengangguran Terhadap Pertumbuhan Ekonomi di Provinsi Jawa Timur Tahun 2016-2018. *IQTISHADIA Jurnal Ekonomi & Perbankan Syariah*, *8*(1), 38–59. https://doi.org/10.19105/iqtishadia.v8i1.4555

Aula, N. K., & Suharto, S. (2021). Pengaruh e-commerce terhadap Produk Domestik Bruto Indonesia. *Jurnal Kebijakan Ekonomi Dan Keuangan*, *1*(1), 39–48. https://doi.org/10.20885/jkek.vol1.iss1.art4

Awalluddin, A. ., & Taufik, I. (2017). Analisis Cluster Data Longitudinal pada Pengelompokan Daerah Berdasarkan Indikator IPM di Jawa Barat. *Prosiding Seminar Nasional Metode Kuantitatif*, *978*, 187–194. https://jurnal.fmipa.unila.ac.id/snmk/article/view/2077/1518

Benri, M., Metisen, H., & Latipa, S. (2015). Analisis Clustering Menggunakan Metode K-Means Dalam Pengelompokkan Penjualan Produk Pada Swalayan Fadhila. *Jurnal Media Infotama*, *11*(2), 110–118. https://core.ac.uk/download/pdf/287160954.pdf

BPS. (2019). Statistik E-Commerce 2019. In *Badan Pusat Statistik*. https://www.bps.go.id/publication/2019/12/18/fd1e96b05342e479a83917c6/statistik-e-commerce-2019.html

Dianary, R. G. F. (2018). Pengaruh E-commerce terhadap Pertumbuhan Ekonomi Indonesia. *Bina Ekonomi*, *22*(1), 43–62.

Fernández-Portillo, A., Almodóvar-González, M., & Hernández-Mogollón, R. (2020). Impact of ICT development on economic growth. A study of OECD European Union countries. *Technology in Society*, *63*(April). https://doi.org/10.1016/j.techsoc.2020.101420

Giswandhani, M., & Hilmi, A. Z. (2020). Pengaruh Kemudahan Transaksi Non-Tunai Terhadap Sikap Konsumtif Masyarakat Kota Makassar. *Jurnal Ilmu Komunikasi*, *9*(2), 239–250.

Haji, K. (2021). E-commerce development in rural and remote areas of BRICS countries. *Journal of Integrative Agriculture*, *20*(4), 979–997. https://doi.org/10.1016/S2095-3119(20)63451-7

Hawari, R., & Kartiasih, F. (2017). Kajian Aktivitas Ekonomi Luar Negeri Indonesia Terhadap Pertumbuhan Ekonomi Indonesia Periode 1998-2014. *Media Statistika*, *9*(2), 119. https://doi.org/10.14710/medstat.9.2.119-132

Jamaludin, K. R. (2022). Dampak Ketersediaan Kurir Online Terhadap Peningkatan Penjualan Usaha Kuliner Di Masamba (Studi Kasus Pelaku Usaha Makan Dan Minum). *Jurnal Ekonomi Pembangunan*.

Karina, D., Alfiyatus Sa'diyah, S., Nabilah, H., & Panorama, M. (2021). Pengaruh Perusahaan Startup Terhadap Pertumbuhan Ekonomi Indonesia Selama Pandemi Covid-19. *Berajah Journal*, *2*(1), 156–166. https://doi.org/10.47353/bj.v2i1.69

Kartiasih, F. (2019a). Dampak Infrastruktur Transportasi Terhadap Pertumbuhan Ekonomi di Indonesia Menggunakan Regresi Data Panel. *Jurnal Ilmiah Ekonomi Dan Bisnis*, *16*(1), 67–77. https://doi.org/10.31849/jieb.v16i1.2306

Kartiasih, F. (2019b). Transformasi Struktural dan Ketimpangan Antardaerah di Provinsi Kalimantan Timur. *INOVASI: Jurnal Ekonomi, Keuangan Dan Manajemen*, *15*(1), 105–113. https://doi.org/https://doi.org/10.30872/jinv.v15i1.5201

Kartiasih, F., Djalal Nachrowi, N., Wisana, I. D. G. K., & Handayani, D. (2023). Inequalities of Indonesia's regional digital development and its association with socioeconomic characteristics: a spatial and multivariate analysis. *Information Technology for Development*, *29*(2–3), 299–328. https://doi.org/10.1080/02681102.2022.2110556

Kartiasih, F., Nachrowi, N. D., Wisana, I. D. G. K., & Handayani, D. (2023a). *Potret Ketimpangan Digital dan Distribusi Pendapatan di Indonesia: Pendekatan Regional Digital Development Index* (1st ed.). UI Publishing.

Kartiasih, F., Nachrowi, N. D., Wisana, I. D. G. K., & Handayani, D. (2023b). Towards the quest to reduce income inequality in Indonesia: Is there a synergy between ICT and the informal sector? *Cogent Economics & Finance*, *11*(2), 1–18. https://doi.org/10.1080/23322039.2023.2241771

Kartiasih, F., & Setiawan, A. (2020). Aplikasi Error Correction Mechanism dalam Analisis Dampak Pertumbuhan Ekonomi, Konsumsi Energi dan Perdagangan Internasional Terhadap Emisi CO2 di Indonesia. *Media Statistika*, *13*(1), 104–115. https://doi.org/10.14710/medstat.13.1.104-115

Khotimah, D. R. K. (2021). Economic Growth and E-Commerce: Potential for Digitizing SMSEs in East Java. *East Java Economic Journal*, *5*(2), 183–203. https://doi.org/10.53572/ejavec.v5i2.69

Kusumasari, A., & Kartiasih, F. (2017). Aglomerasi Industri dan Pengaruhnya Terhadap Pertumbuhan Ekonomi Jawa Barat 2010-2014. *Jurnal Aplikasi Statistika & Komputasi Statistik*, *9*(2), 28–41. https://doi.org/https://doi.org/10.34123/jurnalasks.v9i2.143

Mahir, P. (2015). Klasifikasi Jenis-Jenis Bisnis E-Commerce. *Jurnal Neo-Bis*, *9*(2), 32–40. https://doi.org/10.21107/nbs.v9i2.1271%0A

Mashalah, H. Al, Hassini, E., Gunasekaran, A., & Bhatt (Mishra), D. (2022). The impact of digital transformation on supply chains through e-commerce: Literature review and a conceptual framework. *Transportation Research Part E: Logistics and Transportation Review*, *165*(August), 102837. https://doi.org/10.1016/j.tre.2022.102837

Maulana, A., Arjun, N. R., Akbar, F., Suryanti Ayu, N., & Firmansyah, H. (2021). Peran E-Commerce di Tengah Pandemi terhadap Gaya Hidup Masyarakat Indonesia Masa Kini. *Journal of Education and Technology*, *1*(1), 55–61. http://jurnalilmiah.org/journal/index.php/jet

Muñoz-Villamizar, A., Velázquez-Martínez, J. C., Haro, P., Ferrer, A., & Mariño, R. (2021). The environmental impact of fast shipping e-commerce in inbound logistics operations: A case study in Mexico. *Journal of Cleaner Production*, *283*. https://doi.org/10.1016/j.jclepro.2020.125400

Nasution, E. Y., Hariani, P., Hasibuan, L. S., & Pradita, W. (2020). Perkembangan Transaksi Bisnis E-Commerce terhadap Pertumbuhan Ekonomi di Indonesia. *Jesya*, *3*(2), 506–519. https://doi.org/10.36778/jesya.v3i2.227

Ningsih, Y. P., & Kartiasih, F. (2019). Dampak Guncangan Pertumbuhan Ekonomi Mitra Dagang Utama terhadap Indikator Makroekonomi Indonesia. *Jurnal Ilmiah Ekonomi Dan Bisnis*, *16*(1), 78–92. https://doi.org/https://doi.org/10.31849/jieb.v16i1.2307

Nizar, C., Hamzah, A., & Syahnur, S. (2013). Pengaruh Investasi Dan Tenaga Kerja Terhadap Pertumbuhan Ekonomi Serta Hubungannya Terhadap Tingkat Kemiskinan Di Indonesia. *Jurnal Ilmu Ekonomi*, *1*(2), 1–8.

Nugraha, D. D. C., Naimah, Z., Fahmi, M., & Setiani, N. (2014). Klasterisasi Judul Buku dengan Menggunakan Metode K-Means. *Seminar Nasional Aplikasi Teknologi Informasi (SNATI) Yogyakarta*, *21*(1), 1907–5022.

Pantelimon, F.-V., Georgescu, T. M., & Posedaru, B.-S. (2020). The Impact of Mobile e-Commerce on GDP: A Comparative Analysis between Romania and Germany and How Covid-19 Influences the e-Commerce Activity Worldwide. *Informatica Economica*, *24*(2/2020), 27–41. https://doi.org/10.24818/issn14531305/24.2.2020.03

Pujianto, A., Mulyati, A., & Novaria, R. (2018). Pemanfaatan Big Data Dan Perlindungan Privasi Konsumen di Era Ekonomi Digital. *Majalah Ilmiah Bijak*, *15*(2), 127–137. https://doi.org/10.31334/bijak.v15i2.201

Purnaningrum, E., & Ariqoh, I. (2019). Google Trends Analytics Dalam Bidang Pariwisata. *Majalah Ekonomi*, *24*(2), 232–243. https://doi.org/10.36456/majeko.vol24.no2.a2069

Putri, A. S., & Zakaria, R. (2020). Analisis Pemetaan E-Commerce Terbesar di Indonesia Berdasarkan Model Kekuatan Ekonomi Digital. *Seminar Dan Konferensi Nasional IDEC*, *November*, 1–14.

Siregar, A. M. (2019). Pengelompokan Bidang Laju Pertumbuhan Ekonomi Indonesia Menggunakan Algoritma K-Means. *Jurnal Accounting Information System (AIMS)*, *2*(2), 140–151. https://doi.org/10.32627/aims.v2i2.342

Subian, A. R., Mulkan, D. A., Ahmady, H. H., & Kartiasih, F. (2024). Comparison Methods of Machine Learning and Deep Learning to Forecast The GDP of Indonesia. *SISTEMASI: Jurnal Sistem Informasi*, *13*(1), 149–166. https://doi.org/10.32520/stmsi.v13i1.3445

Totleben, B., & Kuźmar, S. (2022). *Big Data Applications in Economics: Is Google Trends Data Really Useful in Economic Nowcasting? Ekonomiczne i polityczne uwarunkowania upadłości państw View project The use of Big Data in economic nowcasting and forecasting View project Big Data Applicatio*. *November*. https://doi.org/10.13140/RG.2.2.13000.96007

Van Loon, P., Deketele, L., Dewaele, J., McKinnon, A., & Rutherford, C. (2015). A comparative analysis of carbon emissions from online retailing of fast moving consumer goods. *Journal of Cleaner Production*, *106*(2015), 478–486. https://doi.org/10.1016/j.jclepro.2014.06.060

Yulianti, D. T., Damayanti, D., & Prastowo, A. T. (2021). Pengembangan Digitalisasi Perawatan Kesehatan Pada Klink Pratama Sumber Mitra Bandar Lampung. *Jurnal Teknologi Dan Sistem Informasi (JTSI)*, *2*(2), 32–39. http://jim.teknokrat.ac.id/index.php/JTSI