International Journal of Engineering and Computer Science Applications (IJECSA)
Vol. 3, No. 1, March 2024, pp. 43∼50
ISSN: 2828-5611
DOI: 10.30812/IJECSA.v3i1.3890

Clustering Biplot on Tourist Visits in Indonesia

Isma Muthahharah, Zakiyah Mar'ah
Universitas Negeri Makassar, Makassar, Indonesia

Article Info

Article history:

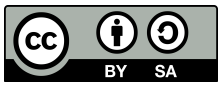
Received February 20, 2024
Revised February 23, 2024
Accepted March 1, 2024

Keywords:

Clustering
K-Medoids
Biplot
Tourist

ABSTRACT

This research aims to find out whether or not many tourists visited Indonesia after Covid-19 by clustering. This will generate foreign exchange earnings and contribute directly to the country's economic growth. **The analytical method used in this research** is K-Medoids. K-Medoids is a partition clustering technique that groups a collection of n objects into k clusters by utilizing the objects in the collection of objects to represent a cluster called a medoid. The data in this research used secondary data related to foreign tourist visits to Indonesia from several publication sources in 2017-2021. **The results of this research show** that there were 3 clusters obtained: Cluster 1 shows the number of tourist visits visiting Indonesia in 2017, 2018 and early 2019 because the Covid-19 pandemic has not yet occurred, Cluster 2 shows that there were no tourist visits in 2020 due to the start of the Covid-19 pandemic, and Cluster 3 indicates low tourist arrivals in 2021 due to the Covid 19 pandemic which temporarily prohibited foreign tourists from visiting Indonesia.



Corresponding Author:

Zakiyah Mar'ah,
Universitas Negeri Makassar, Makassar, Indonesia.
Email: zakiyahm@unm.ac.id

How to Cite: I. Muthahharah and Z. Mar'ah, Clustering Biplot on Tourist Visits in Indonesia, *International Journal of Engineering and Computer Science Applications (IJECSA)*, vol. 3, no. 1, pp. 43-50, Mar. 2024. doi: 10.30812/ijecsa.v3i1.3890.

1. INTRODUCTION

The tourism industry is one of the foreign exchange earning sectors highest for Indonesia. Based on data from the Ministry of Tourism total Indonesia's tourism foreign exchange in 2018 2018 was $19.29M. Besides that based on data from World Travel & Student Council, Indonesia is number one in Southeast Asia as fastest growing country intourism sector [1]. A foreign tourist is any visitor who visits a country other than his/her place of residence for one or more reasons and whose stay is no more than 12 months, without the intention of earning income in the destination country. As the number of foreign tourists increases, the impact will be greater. The positive impact of tourism is that it creates new jobs for local people, increases economic income and makes it possible to learn foreign languages, for example. English, Japanese [2]. There are several factors that influence the number of foreign tourist visits, namely the condition of natural and cultural wealth as well as security issues in terms of environmental health. At the end of 2019, there was a virus outbreak in Wuhan, China which had an impact on human health. Chinese authorities confirmed that the virus was a new corona virus and WHO called this virus COVID-1 which resulted in a lack of tourist visits in Indonesia [3]. The Central Statistics Agency (BPS) recorded that the number of foreign tourist visits to Indonesia during 2022 reached 5,47 million visits. This figure increased by 251.28% compared to the 2021 period. Meanwhile, the number of foreign tourist visits increased from 5,47 million in 2022 to 4,15 million through the main gate, and 1,32 million entered through the border gate. The number of foreign tourist visits is estimated to be the highest in 2022 compared to the COVID-19 pandemic period, namely 2020-2022. If we look at previous data, the number of foreign tourist visitors in 2020 reached 4,05 million people, but in 2021 it will still be at 1,56 million people. Before the corona virus pandemic, the number of foreign tourists in 2019 was 16,11 million people and will increase in 2022.

One analysis process that can be based on distribution patterns in certain clusters with the technique of forming certain cluster groups is Clustering [2]. Clustering is a technique of multivariate analysis which aims to group objects by combining objects in a group of objects that are similar (or related) to each other and different (not related) to objects in other groups [3]. K-Medoid a partition clustering technique that groups a collection of n objects into k clusters by utilizing objects in the collection of objects to represent a cluster called a medoid [4]. Biplot analysis is an attempt to explain data in the form of tables and two-dimensional graphic summaries. With a presentation like this, the characteristics of the variables and objects of observation as well as the relative position between the objects of observation and the variables can be analyzed [5]. Biplot information source contains objects and variables into one photo. Biplot analysis can be expressed visually in a set of objects and variables in a diagram [6]. There are three important things to take away from this Biplots are: 1) Proximity between objects. which objects have similar properties to certain objects, 2) Diversity of variables: This information is used to find out whether there are certain variables whose values are more or less the same for each object, or vice versa, 3) Correlation between variables. This information can be used to assess how a variable influences or is influenced by other variables.

Based on research conducted by [7] it shows that the clustering results of countries with very high, high, quite high, low and very low tourist arrivals use K-Means and PCA modeling. Meanwhile, according to [8] using the K-Mean algorithm obtained results that almost 90% of provinces in Indonesia still have low tourism potential. This is proven by the number of foreign guests staying at hotels based on reports from the central statistics agency. Meanwhile [9] stated that Applying k-means algorithm methods to cluster country of origin, the results showed they were three clusters formed based on the attributes of visiting, i.e. length of stay in Indonesia and total amount of their expenditures. Each cluster consists of 14, 54 and 18 countries. The first cluster is characterized by countries that have high tourism spending; the second cluster is formed by countries with moderate tourism spending; and the third cluster is characterized by countries with low tourism spending. The accuracy of the three clusters in explaining the variance of tourist spending is 68.8 %.

Several relevant studies have not yet grouped tourist visits using biplot clustering, only using Algorithm K-Means and PCA. Therefore, **research aims to** find out whether or not many foreign tourist visits in Indonesia according to the Rupiah exchange rate, the gross regional income, and the population of Indonesia, using cluster and biplot analysis. It is hoped that this research **can contribute** to the Indonesian government in improving tourist destination facilities, supporting training to improve services and performance, and supporting the socialization of tourist attractions, therefore more foreign tourists will visit Indonesia as well as equalizing tourist attractions in the Indonesian region in terms of their development.

2. RESEARCH METHOD

2.1. Data Source

The data in this research uses secondary data related to foreign tourist visits to Indonesia from several publication sources in 2017-2021. The data in this research were taken from four places where the number of Indonesian tourist visits was sourced from

BPS of the Republic of Indonesia, the Rupiah exchange rate was sourced from Bank Indonesia and https://fx-rate.net/historical, GDP per capita was sourced from the World Bank via data.worldbank .org, and the Indonesian population is sourced from Work Bank via data.worldbank.org.

2.2. Research Variables

The variables in this research are as follows:

X_1 : Number of tourist visits to Indonesia
X_2 : Rupiah exchange rate
X_3 : Indonesias GDP per capita
X_4 : Total population of Indonesia

2.3. Research Procedure

The steps taken in this research include:

a) Collect data based on predetermined variables
b) Carry out descriptive statistical analysis tests
c) Carry out cluster analysis by showing the midpoint of the data as well as the year and month with Euclidian distance on Equation (1) [10].

$$d_{ij} = \sqrt{\sum_{n=1}^{p} (q_i - p_i)^2} \tag{1}$$

With, d_{ij} is distance between objects i and j, x_{ia} is value of object i when variable to a, x_{ja} is value of object j when variable to a, and p is Total observed variables. Algorithm for performing K-Medoids clustering (See in Figure 1 [11]

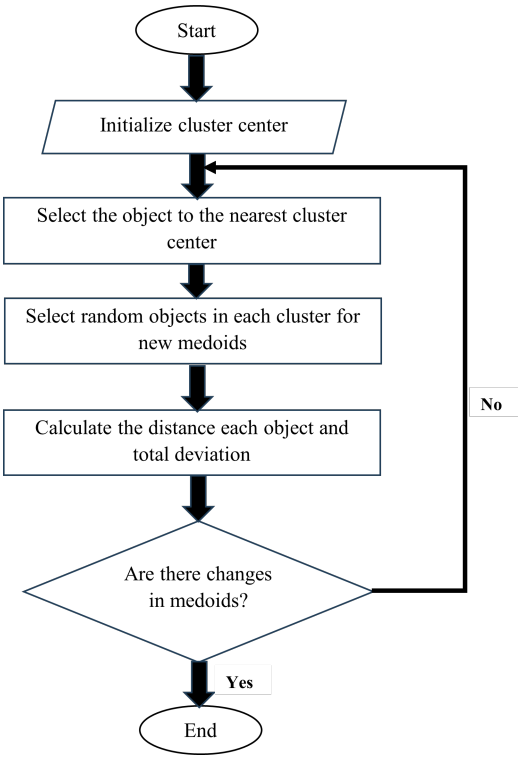


Figure 1. K-Medoids Algorithm

d) Carry out biplot analysis to find out data points that have the same characteristics and correlation between variables using the biplot feasibility measure on Equation (2) [12].

$$p^2 = \frac{\lambda_1 + \lambda_1}{\sum_{k=1}^{r} \lambda_k} \tag{2}$$

Where, p^2 is measure of the feasibility of a two-dimensional biplot for the corresponding value, λ_1 is first largest eigenvalue, λ_2 is second largest eigenvalue, and λ_k is k-th largest eigenvalue, where k = 1,2,3,..,r .

3. RESULT AND ANALYSIS

3.1. Description Analysis

Disciption analysis to determine the number of foreign tourist visits in Indonesia based on the number of Indonesian tourist visits, Rupiah exchange rate, per capita income and the number of Indonesian residents for 5 years with the total population of Indonesia. To see how large the number of foreign tourist visits is, see Figure 2.

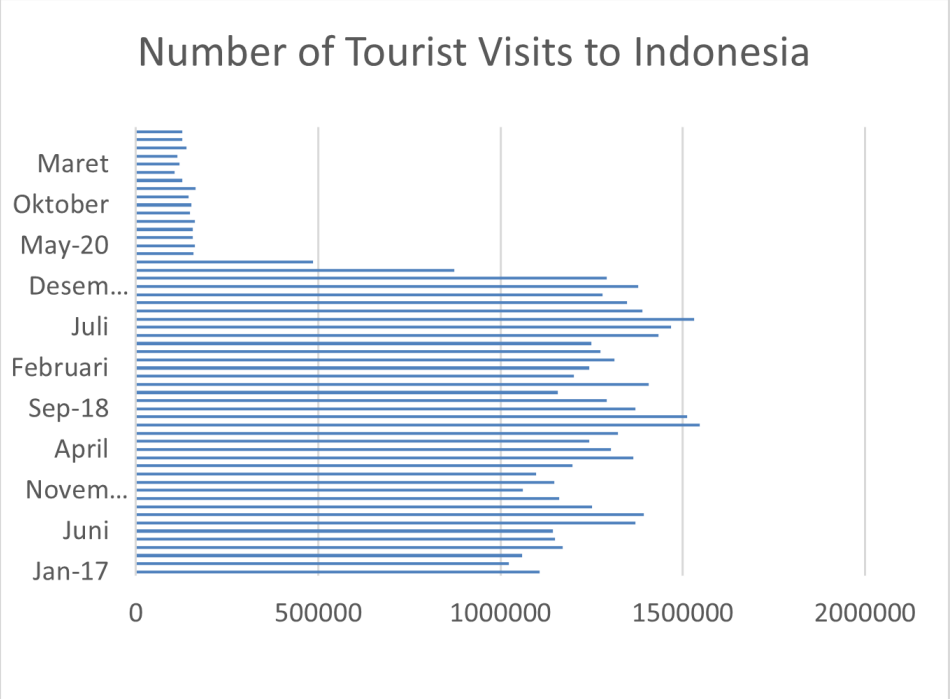


Figure 2. Number of Tourist Visits to Indonesia

Figure 2 shows that the highest number of foreign tourist visits was in September 2018 and September 2021, namely around 1,500,000 people, while the lowest number in May 2020 and May 2012 did not reach 500,000 people. To find out how much the rupiah exchange rate occurred during visits by foreign tourists to Indonesia from 2017 to 2021, you can see in Figure 3. From Figure 3 it is known that the highest rupiah to USD exchange rate occurred in 2021, namely 14,900, while the lowest rupiah to USD exchange rate occurred in 2018, namely 14,000. To find out gross regional income in 2017-2021, see Figure 4.

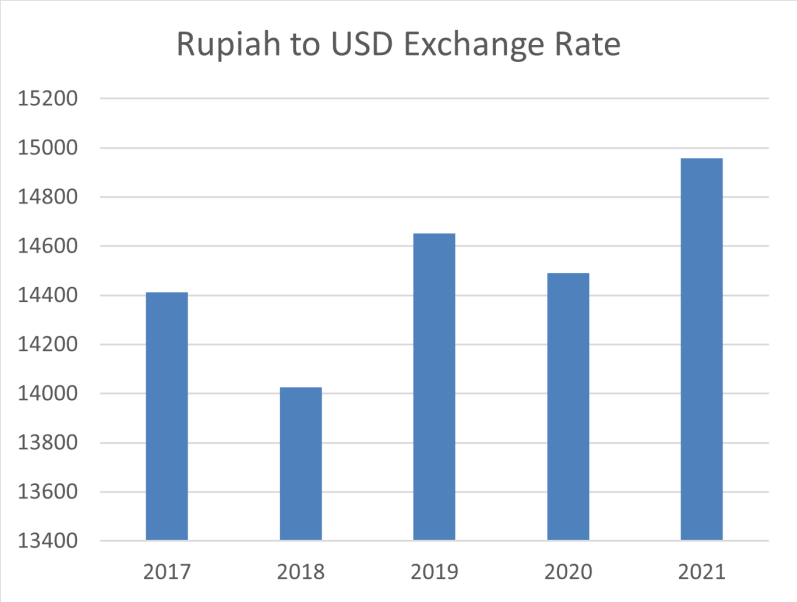


Figure 3. Rupiah to USD Exchange Rate

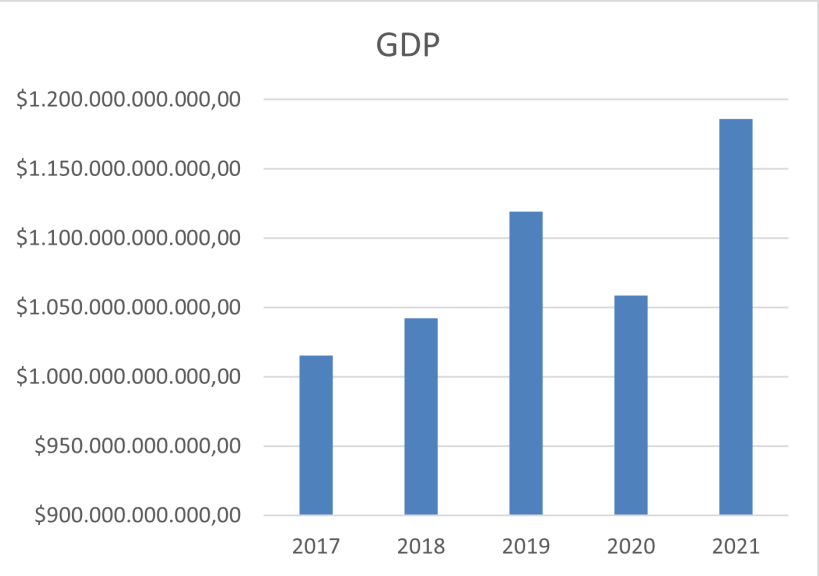


Figure 4. GDP

Figure 4 shows that the largest gross regional income in 2021 is $1,190,000,000,000.00 and the smallest is in 2017, namely $1,020,000,000,000.00.

3.2. K-Medoid Clustering

K-Medoid are used to collect objects that are very far away in location/characteristics so that if they are inserted into a cluster the data can distort the average (mean) value of the cluster. K-medoid is used to group foreign tourist visits based on the number of tourist visits to Indonesia, the Rupiah exchange rate, Indonesian GDP per capita and the number of Indonesian residents by looking at the midpoint or collection of objects that represent a cluster (See in Figure 5).

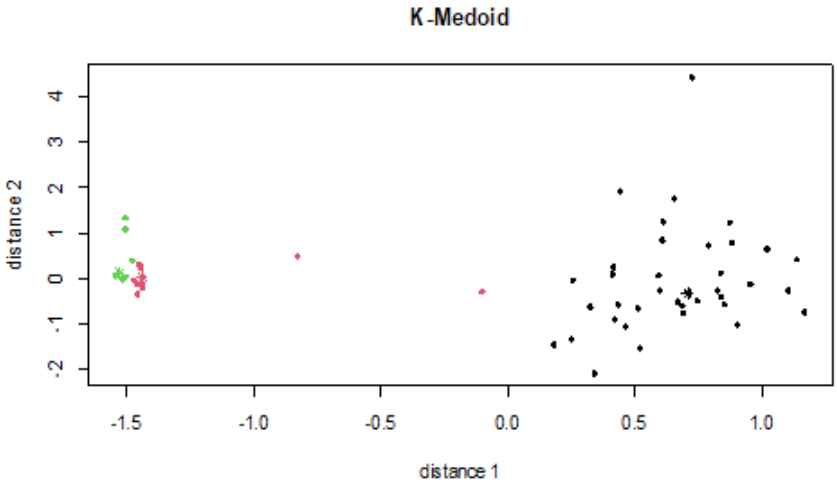


Figure 5. K-Medoid

From Figure 5, the findings of this research are known that some data distributions are located at a distance of -1.5, one data is at a distance of -0.8 and is close to 0.0 (this is the midpoint of the data distribution). Other data is between 0.0-1.2. To find out the distribution of data by month and year, see Figure 6.

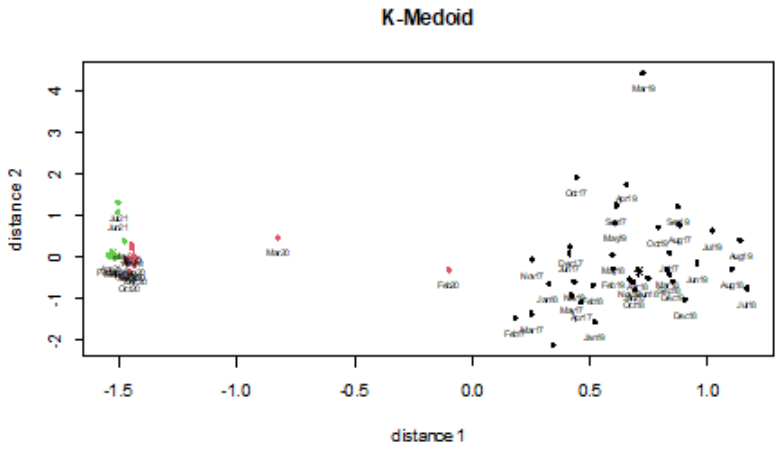


Figure 6. K-Medoids (Month and Year)

From Figure 6, the findings of this research are a distance of -1.5 there are several points/data in February-December 2020 and May, June, July 2021, a distance of -0.8 in March 2020, and a distance approaching 0.0 in February 2020. Meanwhile, the data at 0.0-1.2 was in January-December 2017, 2018 and 2019.

3.3. Clustering Biplot

Biplot analysis is a descriptive statistical tool that simultaneously represents the relative position of objects observed in two dimensions with the variable p. This analysis is based on singular value decomposition (PNS) of the matrix and is applied to data with an interval or ratio measurement scale and can also be used for ordinal data. To find out the dimensions of the point/data distribution, see Figure 7.

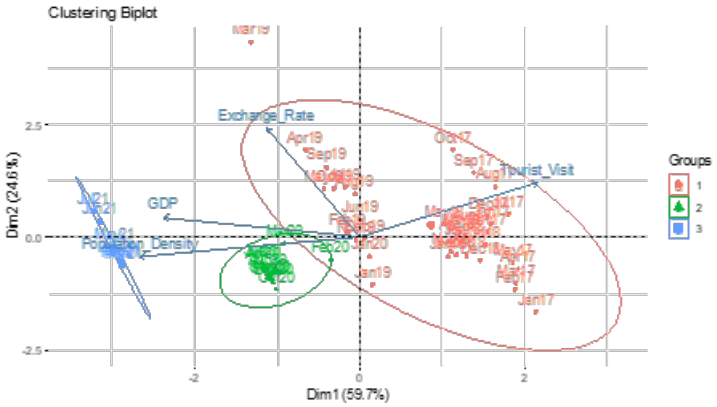


Figure 7. FP-Growth Method with RapidMiner

From Figure 7, **the findings of this research** are points in the same group have the same characteristics. Variables that have small variations are GDP and population density variables. Meanwhile, the variables that have great diversity are the exchange rate and tourist visits variables. The correlation between variables between GDP and population density shows a strong positive correlation, meaning that the greater the population, the higher the GDP. Meanwhile, population density and tourist visits show a weak negative correlation, meaning that the higher the population, the lower the number of foreign tourist visits. Many foreign tourists visited Indonesia in 2017, 2018 and early 2019, this is because in 2019 there was no Covid-19 pandemic. 2019 and early 2020 had strong characteristics in the exchange rate variable, this is because that year was the beginning of Covid-19. 19 so that the money exchange rate continues to rise. The years 2020 and early 2021 have strong characteristics in the GDP and population density variables, this is because that year was a period of lockdown which required staying at home to avoid the spread of Covid-19

The results of this research are **in line with or supported** by [13] which states that there were significant differences in the tourism sector before COVID-19 and when COVID-19 occurred. This occurred as a result of travel restrictions imposed by the government to reduce the spread of COVID-19, resulting in a reduction in both domestic and international travel. The spread of the COVID-19 pandemic has caused foreign tourists to cancel trips and bookings on a large scale and the reluctance of tourists to visit is influenced by concerns about the impact of tourism travel during the pandemic. Meanwhile, research [14] states that Decline in foreign and domestic tourism as well due to the temporary ban on foreign tourists entering Indonesia and restrictions domestic tourists go to several cities and provinces in Indonesia, mode restrictions air and sea transportation. There is an influence of the Covid-19 pandemic on hotel occupancy rates in Indonesia. The longer the Covid 19 pandemic lasts, the more the occupancy rate for star hotels in Indonesia is increasingly decreasing due to restrictions activities in public places or facilities, restrictions on modes of transportation, hotel room prices, quality of hotel services and hotel facilities.

4. CONCLUSION

From the research it can be concluded that 1) clusters with red dots indicate the number of Tourist Visits who visited Indonesia in 2017, 2018 and early 2019 because the Covid-19 pandemic has not yet occurred, 2) clusters with green dots in 2020 indicate that Tourist Visits are not very many due to the beginning of the Covid-19 pandemic. 3) Clusters with blue dots indicate that the lack of Tourist Visits in 2021 was due to the Covid-19 pandemic which temporarily prohibited foreign tourists from visiting Indonesia. In this case, researchers suggest carrying out mapping using different methods and variables with the aim of perfecting research on foreign tourist visits to Indonesia.

5. DECLARATIONS

AUTHOR CONTIBUTION

All authors contributed to the writing of this article

FUNDING STATEMENT

This research received no financial support from any party.

COMPETING INTEREST

The authors declare no competing interests regarding the data presented in this research.

REFERENCES

[1] A. A. Sudjana, S. N. Aini, and H. K. Nizar, "Revenge Tourism: Analisis Minat Wisatawan Pasca Pandemi Covid-19," *Pringgitan*, vol. 2, no. 01, pp. 1–10, mar 2021. [Online]. Available: http://ejournal.stipram.ac.id/index.php/pringgitan/article/view/158

[2] P. Nurul Sabrina, A. Kania Ningsih, and F. Kasyidi, "Interpretasi dan Visualisasi Hasil Clustering Menggunakan K-Medoid untuk Identifikasi Penyebaran Virus Covid-19," *Jurnal ICT : Information Communication and Technology*, vol. 21, no. 1, pp. 1–7, 2022.

[3] I. Muthahharah and A. Juhari, "A Cluster Analysis with Complete Linkage and Ward's Method for Health Service Data in Makassar City," *Jurnal Varian*, vol. 4, no. 2, pp. 109–116, 2021.

[4] M. A. Nahdliyah, T. Widiharih, and A. Prahutama, "Metode K-Medoids Clustering dengan Validasi Silhouette Index dan C-Index," *Jurnal Gaussian*, vol. 8, no. 2, pp. 161–170, 2019.

[5] R. O. M. Pogalin, C. E. Mongi, and N. Nainggolan, "Analisis Biplot Untuk Pemetaan Kabupaten/Kota di Provinsi Sulawesi Utara Berdasarkan Beberapa Variabel Pendidikan," *Jurnal MIPA*, vol. 10, no. 1, pp. 1–4, sep 2020. [Online]. Available: https://ejournal.unsrat.ac.id/index.php/jmuo/article/view/30416

[6] D. Londong Padang, A. J. Rindengan, and J. S. Kekenusa, "Analisis Biplot pada Pemetaan Karakteristik Kemiskinan di Provinsi Sulawesi Utara," *d'CARTESIAN*, vol. 8, no. 1, pp. 49–52, mar 2019. [Online]. Available: https://ejournal.unsrat.ac.id/index.php/decartesian/article/view/24595

[7] E. Muningsih, N. Hasan, and G. B. Sulistyo, "Penerapan Metode Principle Component Analysis (PCA) untuk Clustering Data Kunjungan Wisatawan Mancanegara ke Indonesia," *Bianglala Informatika*, vol. 8, no. 1, pp. 58–62, 2020.

[8] R. Wulan Sari and D. Hartama, "Data Mining: Algoritma K-Means Pada Pengelompokkan Wisata Asing ke Indonesia Menurut Provinsi," in *Seminar Nasional Sains and Teknologi Informasi (SENSASI)*, 2018, pp. 322–326.

[9] A. A. Mhardika, E. N. Kencana, I. K. G. Sukarsa, K. Jayanegara, I. L. Wijayakusuma, and I. W. Sumarjaya, "Klasterarisasi Karakteristik Wisatawan Manacanegara Menggunakan Metode K-Maenas Clustering," *E-Jurnal Matematika*, vol. 12, no. 2, pp. 140–148, may 2023. [Online]. Available: https://ojs.unud.ac.id/index.php/mtk/article/view/99180

[10] M. S. Pangestu and M. A. Fitriani, "Perbandingan Perhitungan Jarak Euclidean Distance, Manhattan Distance, dan Cosine Similarity dalam Pengelompokan Data Bibit Padi Menggunakan Algoritma K-Means," *Sainteks*, vol. 19, no. 2, pp. 141–155, oct 2022. [Online]. Available: http://jurnalnasional.ump.ac.id/index.php/SAINTEKS/article/view/14495

[11] B. Wira, A. E. Budianto, and A. S. Wiguna, "Implementasi Metode K-Medoids Clustering Untuk Mengetahui Pola Pemilihan Program Studi Mahasiswa Baru Tahun 2018 Di Universitas Kanjuruhan Malang," *RAINSTEK : Jurnal Terapan Sains and Teknologi*, vol. 1, no. 3, pp. 53–68, 2019.

[12] I. Sulistiana and Hidayati, "Pemetaan Karakterstik Kemiskinan Dengan Analisis Biplot Pada Kabupaten/Kota Kepulauan Bangka Belitung," *UNNES Journal of Mathematics*, vol. 9, no. 2, pp. 125 – 130, 2020.

[13] K. Normasyhuri, A. Habibi, and E. Anggraeni, "Fenomena gejolak ekonomi pada sektor perdagangan dan sektor pariwisata: studi perbandingan sebelum dan selama Covid-19," *Jurnal Paradigma Ekonomika*, vol. 17, no. 1, pp. 181–188, 2022.

[14] S. Soehardi, L. Purnamaasih, and D. Rapitasari, "Dampak Pandemik Covid-19 Terhadap Kunjungan Turis Asing dan Domestik serta Tingkat Hunian Kamar Hotel Bintang di Indonesia," *Jurnal Kajian Ilmiah*, vol. 20, no. 3, pp. 291–308, 2020.