Countries' economic segmentation using k-means clustering for the year 2021

Souhila BENBRAHIM [1*], Saloua Nassima CHAOUCHE [2], Rachid TOUMACHE [3]

[1] LEQAD, ENSSEA (Algeria), benbrahim.souhila@enssea.net
[2] LEQAD, ENSSEA (Algeria), chaouche.saloua@hotmail.com
[3] LEQAD, ENSSEA (Algeria), rtoumache@gmail.com

Received: 11/11/2022 **Accepted:** 04/12/2022 **Published:**31/12/2022

Abstract:

The wealth of nations was previously measured by the increase in GDP, but after the inclusion of the social and environmental axes in the measurement of sustainable development, this is no longer the case. This wealth is no longer seen as the sole economic view, but that does not prevent it from remaining the most important axis for measuring the wealth of nations.

The objectives of this study were to: (i) do Value-segmentation of countries for economic axis and (ii) classify them based on collected data patterns and give their characteristics.

In this paper, we have summarized the results of the study that aims to break down countries into several distinct groups using data mining clustering algorithms. This behavioral-based Value-segmentation was done using k-means as a data mining clustering algorithm using GDP, GDP per capita, and GDP per area. This analysis classified countries into five value segments using k-means and selected GDP per capita as the best metric to measure the economic axis.

Keywords: K-means; Segmentation; Data mining; Economic; GDP.

JEL Classification Codes: C61, C88, O50

** Corresponding author.*

1-Introduction :

The growth analysis requires to measure GDP of the economy, in which the GDP formula presents what are driving factors of economic growth (Trinh, 2017).

Clustering is the task of segmenting a heterogeneous population into a number of more homogeneous subgroups or clusters. Clustering differs from classification in that it does not rely on pre-defined categories. The groups in cluster models are not really known in advance. It is up to the algorithms to analyze the input data patterns and identify the natural groupings of instances/cases. When new cases are scored by the generated cluster model, they are assigned to one of the revealed clusters. Clustering algorithms include agglomerative or hierarchical, K-means, K-medoids, Twostep cluster, and Kohonen Network/Self-Organizing Map.

K-Means Clustering is one of the most widely used clustering algorithms. The "K" in the algorithm's name indicates that it searches for a set number of clusters that are defined by the distance between the data points. This algorithm is a straightforward and elegant method for dividing a data set into K distinct, non-overlapping clusters. K-means clustering is carried out by specifying the desired number of clusters K, after which the K-means algorithm assigns each observation to exactly one of the K clusters. It procedure is the result of a simple and straightforward mathematical problem.

This study consists of creating segments of countries using clustering algorithm in order to do value-segmentation of countries for the economic axis using k-means clustering, classify them using collected data patterns and give their characteristics, and select the best metric to segment countries for the economic axis.

The primary research question in this study is: **How can the k-means clustering algorithm classify countries for economic axis?**

In order to answer this question and structure the work in this study, we will be trying to answer a set of questions throughout the document, including:

- What is the best method used for k-means clustering?
- Can k-means clustering iterate good solutions for country segmentation?
- What is the best metric to segment countries for the economic axis?

To answer these questions, we suppose the following hypotheses:

- The best method used for k-means clustering is the one made after the data is cleaned by treating null values, normalizing and standardizing the data, and treating outliers.
- K-means clustering iterates good solutions in country segmentation.
- GDP is the best metric to segment countries for the economic axis.

In order to provide answers to our concerns and to confirm or deny our assumptions, we have organized our paper as follows:

This paper begins with an introduction giving an overview of our work, followed by data preparation focusing on how we get the raw data, identify relevant variables, and do the data cleaning, followed by data treatment, which consists of using K-means clustering for the three metrics as follows: GDP, GDP per capita, and GDP per area, as well as the choice of the best metric, and finally a conclusion that summarizes all the results of our study.

2. Literature review:

Countries segmentation is inspired by customer segmentation which is defined as the process of dividing the customer base into distinct and internally homogeneous groups so that differentiated marketing strategies can be developed based on their unique characteristics (Konstantinova et al., 2020).

In consumer markets, the following are used as criteria (Peelen & Beltman, 2022):

- Geographical characteristics such as postcode;
- Demographic characteristics such as age or sex;
- Socioeconomic data such as income, social class or education;
- Behavior such as purchase or communication behaviour;
- Psychographic characteristics such as lifestyle and the set of norms and values;
- Buying motives and purchase considerations.

Thus, customer segmentation can be distinguished by the following segmentation types (Tsiptsis & Chorianopoulos, 2011):

- **Value based**: In value-based segmentation Customers are divided into groups based on their financial value. This is one of the most important types of segmentation because it can be used to identify the most valuable customers as well as track their value and changes over time. It is also used to distinguish service delivery strategies and optimize resource allocation in marketing campaigns.
- **Behavioral:** This is a very effective and practical segmentation method. It is also widely used because it has few drawbacks in terms of data availability. These data are typically stored and accessible in the organization's databases. Customers are divided into groups based on their observed behavior and usage patterns. This type of segmentation is frequently used to create tailored product offering strategies as well as developing new products and designing loyalty schemes.
- **Propensity based:** Customers are segmented based on their propensity scores, like churn scores, cross-selling scores, which are estimated by classification models. These scores can be combined with other segmentation schemes in order to get better marketing options.
- **Loyalty based:** Loyalty segmentation entails determining the loyalty status of customers and identifying loyalty-based segments such as loyal and switchers/migrators. Cross-selling on prospectively loyal customers can then be focused on high-value customers with a disloyal profile, while retention actions can be focused on high-value customers with a disloyal profile.
- **Socio-demographic and life-stage:** This type identifies various customer groups based on socio-demographic and/or life-stage data such as age, income, and marital status. This type of segmentation is useful for promoting specific life-stage-based products and for assisting with life-stage marketing.
- **Needs/attitudinal:** This type of segmentation is usually based on market research data, and it divides customers into groups based on their needs, desires, attitudes, preferences, and perceptions of the company's services and products. It can be used to assist in the development of new products as well as to determine the brand image and key product features that should be communited.

3. Methodology:

3.1 Getting Raw Data:

To get a better understanding of what type of data should be used for this study, the desired behavior should be addressed. Raw data must describe the problem appropriately and accurately. The raw data used in this study were collected from different sources database as a Comma-Separated Values (CSV) files as follow:

- Country code alpha-3 as CSV file (Github, 2020);
- GDP as CSV file (*GDP (current US$) | Data*, 2022);

- POPULATION as CSV file (*Population, total | Data*, 2022);
- AREA as CSV file (*Land area (sq. km) | Data*, 2022).

The two-dimensional data table from which the data were extracted consists of four columns and 192 rows, it contains annual countries data in 2021. The columns in the table are considered AS descriptive variables and the rows as countries. The choice of the year 2021 is referred to the availability of data.

The following table describes the different used variables:

Table 1. Description of the Raw Variables Used in the Study

Variable name	Variable type	Description
Country Code	STRING	The country's alpha-3 code as described in the ISO 3166 international standard (*ISO - ISO 3166 — Country Codes*, s. d.).
GDP	INTEGER	Monetary measure of the market value of all the final goods and services produced in a period of time in a country (Agarwal, 2018).
POPULATION	INTEGER	The whole number of people in a country.
AREA	INTEGER	The surface of the country.

Source: Done by authors

3.2 Identifying Relevant Variables:

The relevant variables in our case are GDP, GDP_POP, GDP_AREA. GDP_POP is calculated through dividing GDP by POPULATION which represent GDP Per Capita and GDP_AREA is calculated through dividing GDP by AREA. Those two variables are created in order to adopt a more normalized approach.

3.3 Data Cleaning:

The goal of data cleaning is to ensure there are no errors (or as few as possible) that could influence our analysis, it is ensured by order as follows:

3.3.1. Treating NaN values:

In our case, we don't have null values because they are treated in data collection.

3.3.2. Standardizing variables:

Transform relevant variables to have a mean of 0 and a standard deviation of 1.

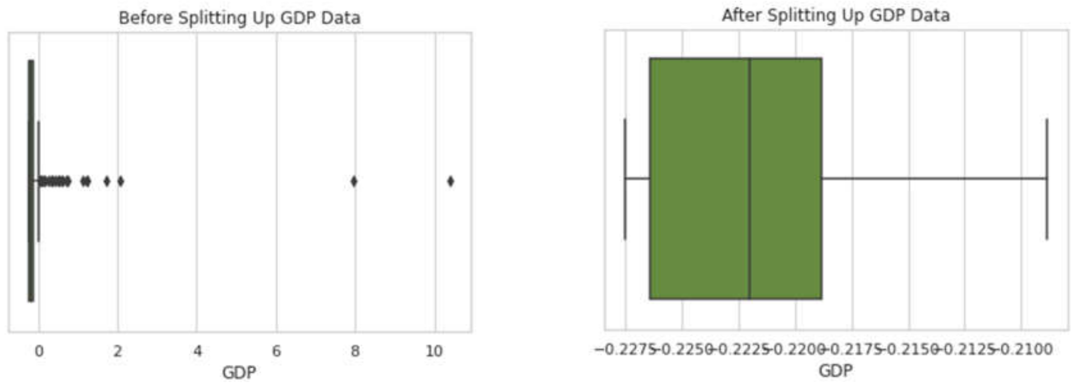
3.3.3. Detecting and treating Outliers:

The K-means clustering algorithm is sensitive to outliers due to the influence of the mean by extreme values. To address this problem, a univariate analysis has been conducted on each relevant variable in order to detect outliers. The purpose of outliers' detection here is not to permanently remove them but to only remove them from the first k-means clustering processing to get more significant limits of clusters' ranges.

3.3.3.1. GDP:

The following figure compares the distribution of Standardized GDP using boxplots before and after splitting up GDP data:

Fig.1. Boxplots of GDP Data Before and After Splitting Them Up



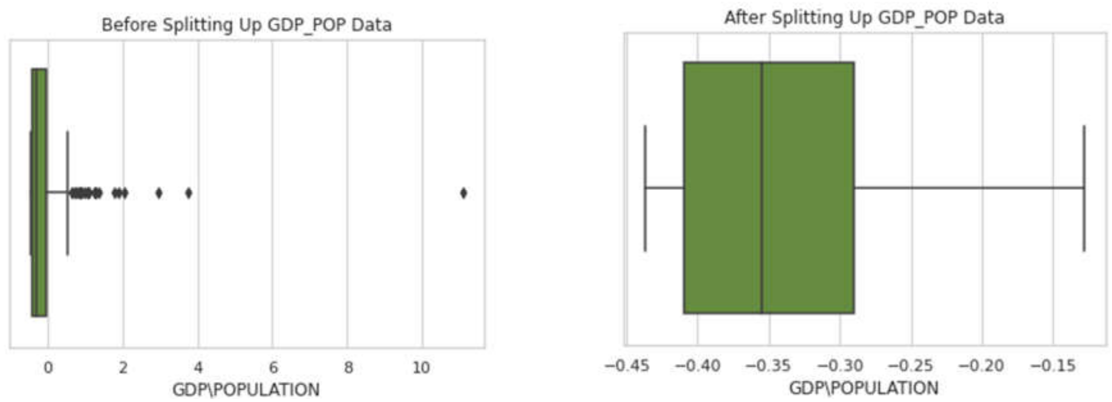
Source: Done by authors using Python

Fig.1. shows that there is a large number of outliers in GDP data before splitting them up. Outliers are defined as the data point located outside the box plot's whiskers. After removing them, the distribution of the data was much better.

3.3.3.2.GDP Per Capita:

The following figure compares the distribution of GDP Per Capita using boxplots before and after splitting up GDP_POP data:

Fig.2. Boxplots of GDP_POP Data Before and After Splitting Them Up



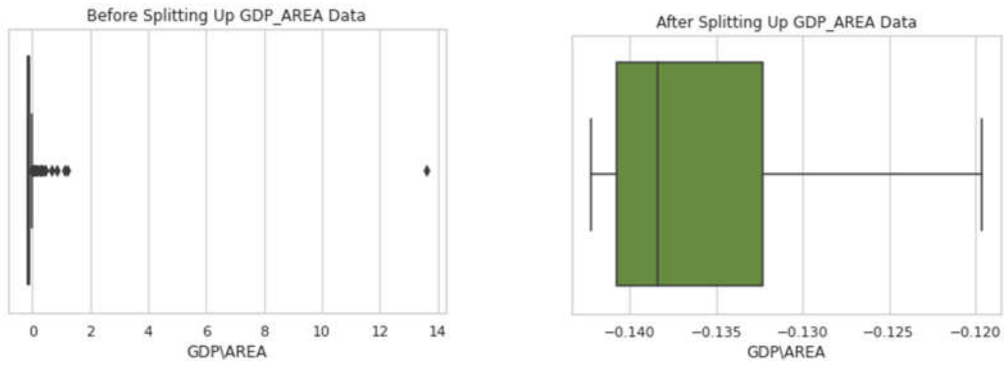
Source: Done by authors using Python

Fig.2. shows that there is a large number of outliers in GDP_POP data before splitting them up.

3.3.3.3.GDP Per AREA:

The following figure compares the distribution of GDP Per Area using boxplots before and after splitting up GDP_AREA data:

Fig.3. Boxplots of GDP_AREA Data Before and After Splitting Them Up



Source: Done by authors using Python

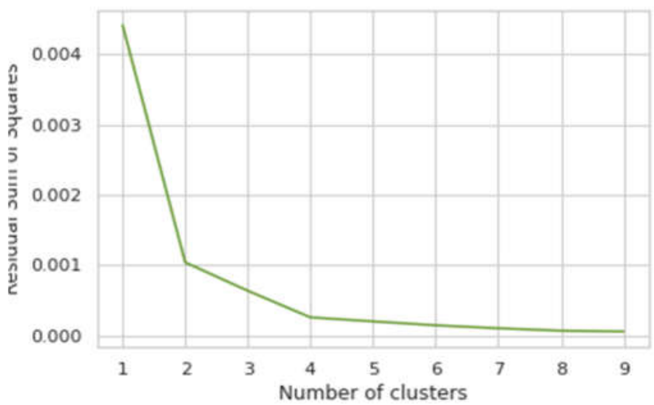
Fig.3. shows that there is a large number of outliers in GDP_AREA data before splitting them up.

4. Results:

4.1 GDP Clustering:

The optimal number of clusters for GDP after splitting the data up was selected using the "elbow method" as shows the following figure:

Fig.4. Selection of the Optimum Number of Clusters for GDP After Splitting Up



Source: Done by authors using Python

The k-means clustering process was applied after splitting up and removing outliers and according to Elbow method, the number of clusters for GDP was set to 3.

The next step is to apply the k-means algorithm to the removed GDP data with two clusters.

The obtained clusters are detailed in the following figure:

Fig.5. The range of different clusters of GDP

GDPCluster	count	mean	std	min	25%	50%	75%	max
0	46.0	-0.226186	0.001188	-0.227511	-0.227089	-0.226609	-0.225466	-0.223504
1	38.0	-0.219974	0.001576	-0.222431	-0.221206	-0.219859	-0.218866	-0.217123
2	16.0	-0.212475	0.002696	-0.215790	-0.214867	-0.212987	-0.209600	-0.208878
3	90.0	0.042435	0.420310	-0.208211	-0.188375	-0.112181	0.001755	2.051335
4	2.0	9.171999	1.717352	7.957648	8.564824	9.171999	9.779175	10.386351

Source: Done by authors using Python

The assigning of GDP segment names is shown in the following table:

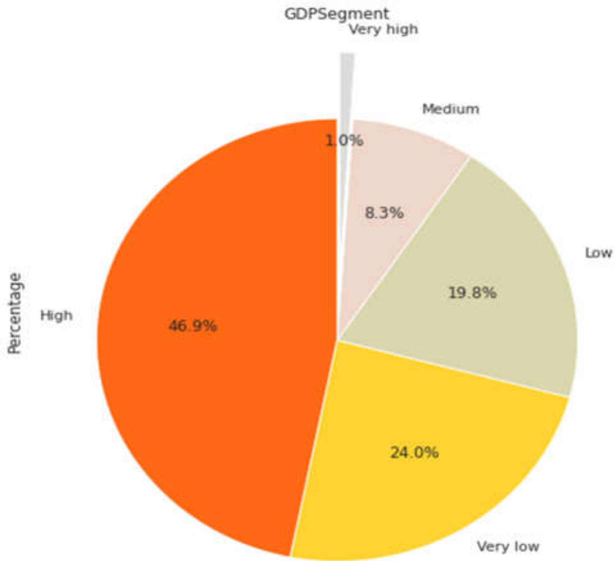
Table 2. Assigning GDP segment names

GDP Cluster Number	GDP Segment
0	Very low
1	Low
2	Medium
3	High
4	Very high

Source: Done by authors

As shown in **Table 2**, the five resulting segments of GDP varying from very low to very high are illustrated in the figure below, which gives the numerical proportion of each segment.

Fig.6. Pie chart of GDP Segment



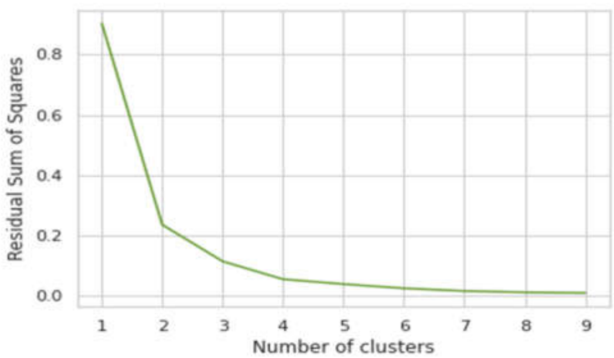
Source: Done by authors using Python

Fig.6. shows that observations that are in the "High" GDP segment make up a proportion of 46.9%, which is the largest proportion of the population.

4.2 GDP_POP Clustering:

The optimal number of clusters for GDP Per Capita after splitting the data up was selected using the "elbow method" as shows the following figure:

Fig.7. Selection of the Optimum Number of Clusters for GDP_POP After Splitting Up



Source: Done by authors using Python

The k-means clustering process was applied after splitting up and removing outliers and according to Elbow method, the number of clusters for GDP Per Capita was set to 3.

The next step is to apply the k-means algorithm to the removed GDP_POP data with two clusters.

The obtained clusters are detailed in the following figure:

Fig.8. Selection of the Optimum Number of Clusters for GDP_POP After Splitting Up

GDP_POPCluster	count	mean	std	min	25%	50%	75%	max
0	63.0	-0.403818	0.039841	-0.436721	-0.424372	-0.412941	-0.391457	-0.128487
1	44.0	-0.325379	0.026480	-0.364397	-0.347512	-0.328785	-0.310109	-0.267651
2	27.0	-0.208556	0.040613	-0.261667	-0.242316	-0.208283	-0.192820	-0.129681
3	56.0	0.545659	0.637513	-0.109712	0.055266	0.330951	0.867287	2.952133
4	2.0	7.415653	5.168338	3.761086	5.588369	7.415653	9.242936	11.070220

Source: Done by authors using Python

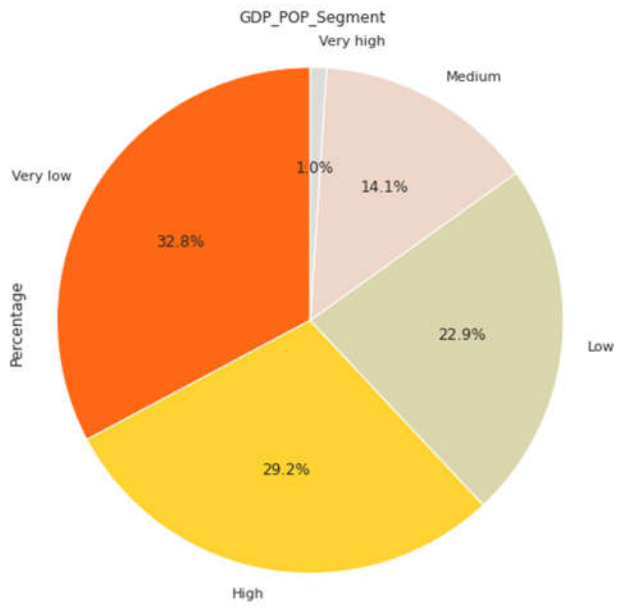
The assigning of GDP_POP segment names is shown in the following table:

Table 3. Assigning GDP_POP segment names

GDP_POP Cluster	GDP_POP Segment
0	Very low
1	Low
2	Medium
3	High
4	Very high

Source: Done by authors

Fig.9. Pie chart of GDP_POP Segment



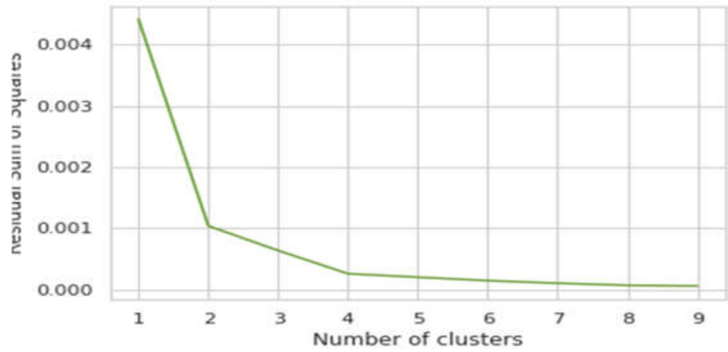
Source: Done by authors using Python

Fig.9. shows that observations that are in the "Very low" GDP_POP segment make up a proportion of 32.8% which is the largest proportion of the population.

4.3 GDP_AREA Clustering:

The optimal number of clusters for GDP Per Area after splitting the data up was selected using the “elbow method” as shows the following figure:

Fig.10. Selection of the Optimum Number of Clusters for GDP_AREA After Splitting Up



Source: Done by authors using Python

The k-means clustering process was applied after splitting up and removing outliers and according to Elbow method, the number of clusters for GDP_AREA was set to 3.

The next step is to apply the k-means algorithm to the removed GDP_AREA data with two clusters.

The obtained clusters are detailed in the following figure:

Fig.11. The range of different clusters of GDP_AREA

GDP_AREACluster	count	mean	std	min	25%	50%	75%	max
0	71.0	-0.140323	0.001462	-0.142246	-0.141595	-0.140613	-0.139303	-0.137033
1	32.0	-0.133629	0.002062	-0.136868	-0.135533	-0.134053	-0.132024	-0.129912
2	21.0	-0.125125	0.002721	-0.129092	-0.127055	-0.124835	-0.122957	-0.119652
3	67.0	0.048976	0.274706	-0.118646	-0.103943	-0.070643	0.058119	1.193204
4	1.0	13.585287	NaN	13.585287	13.585287	13.585287	13.585287	13.585287

Source: Done by authors using Python

The assigning of GDP_AREA segment names is shown in the following table:

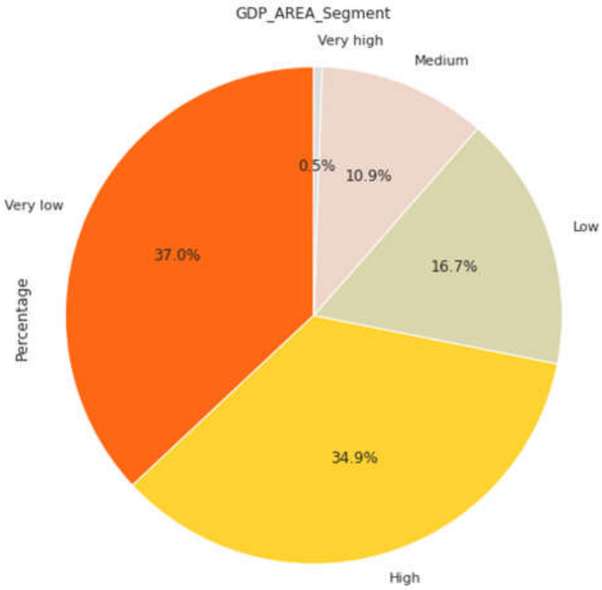
Table 4. Assigning GDP_AREA segment names

GDP_AREA Cluster	GDP_AREA Segment
0	Very low
1	Low
2	Medium
3	High
4	Very high

Source: Done by authors

As shown in **Table 4**, the five resulting segments of GDP Per Area varying from Very low to Very high are illustrated in the figure below, which gives the numerical proportion of each segment.

Fig.12. Pie chart of GDP_AREA Segment



Source: Done by authors using Python

Fig.12. shows that observations that are in the "Very low" GDP segment make up a proportion of 37% which is the largest proportion of the population.

4.4 Best metric:

The best metric, giving results that are more realistic and logical compared to the segmentation of countries presented by the World Bank (*World Bank Open Data | Data*.), is GDP_POP.

The results of the segmentation are presented in table 5 in the appendices, the five resulting segments of GDP Per capita are summarized as bellow:

- Sixty-three countries are classified in the very low segment with the proportion of 32.8%.
- Forty-four countries are classified in the low segment with the proportion of 22.9%.
- Twenty-seven countries are classified in the medium segment with the proportion of 14.1%.
- Fifty-six countries are classified in the high segment with the proportion of 29.2%.
- Two countries are classified in the very high segment with the proportion of 1% which are presented by Bhutany and Liechtenstein.

5-Conclusion:

Throughout our research, we have tried to answer a primary question that relates to finding how the k-means clustering algorithm can classify countries for economic axis. The main objective of our research was to create segments of countries for the economic axis using k-means clustering, classify them using collected data patterns and give their characteristics, and select the best metric to segment countries for the economic axis, against which the following can be concluded:

As k-means is sensitive to extreme values, we were obliged to clean it by only standardizing and treating extreme values by removing them from the first k-means clustering processing and treating them in a second k-means clustering processing.

One of the main findings, in addition to finding that K-means clustering iterates good solutions in country segmentation, is that GDP is not the best metric to segment countries for the economic axis but rather the GDP per capita.

One of the main limitations of the study was the difficulty of obtaining articles that addressed the K-means clustering for macro studies and also the availability of all country data, for those reasons, we limited the sample to only 192 countries.

As for possible lines of research, the space is left open to segment countries along other axes.

6. Bibliography List :

- **Books :**
  - Konstantinova, L. A., Kramarenko, I. V., Denisova, A. I., & Margarov, G. I. (2020). Differentiation of Clients Based on Behavioral Data Using Domestic Software. Institute of Scientific Communications Conference, 1623-1631.
  - Peelen, E., & Beltman, R. (2022). Customer Relationship Management (An). ESCE CURRICULUM, 2, 128.
  - Tsiptsis, K. K., & Chorianopoulos, A. (2011). Data mining techniques in CRM : Inside customer segmentation. John Wiley & Sons.
- **Journal article :**
  - Agarwal, S. C. (2018). Gross Domestic Product & it's Various Approaches. International Journal of Tax Economics and Management, 1(2).
  - Trinh, T. H. (2017). A Primer on GDP and Economic Growth. International Journal of Economic Research, 14(5), 13.
- **Internet websites:**
  - GDP (current US$) | Data. (2022). https://data.worldbank.org/indicator/NY.GDP.MKTP.CD (consulted on 06/09/2022).
  - Github. (2020, mai 18). Countries with their (ISO 3166-1) Alpha-2 code, Alpha-3 code, UN M49, average latitude and longitude coordinates. Gist. https://gist.github.com/tadast/8827699 (consulted on 06/09/2022).
  - ISO - ISO 3166—Country Codes. (s. d.). ISO. https://www.iso.org/iso-3166-country-codes.html (consulted on 06/09/2022).
  - Land area (sq. Km) | Data. (2022). https://data.worldbank.org/indicator/AG.LND.TOTL.K2 (consulted on 06/09/2022).
  - Population, total | Data. (2022). https://data.worldbank.org/indicator/SP.POP.TOTL (consulted on 06/09/2022).
  - World Bank Open Data | Data. (s. d.). https://data.worldbank.org/ (consulted on 06/09/2022).

7. Appendices:

Table 5. Segmentation results

Country Code	Country Name	GDP Cluster	GDP_POP Cluster	GDP_AREACluster	GDP Segment	GDP_POP Segment	GDP_AREA Segment
AFG	Afghanistan	1	0	0	Low	Very low	Very low
AGO	Angola	3	0	0	High	Very low	Very low
ALB	Albania	1	1	2	Low	Low	Medium
AND	Andorra	0	3	3	Very low	High	High

ARE	United Arab Emirates	3	3	3	High	High	High
ARG	Argentina	3	2	0	High	Medium	Very low
ARM	Armenia	1	1	1	Low	Low	Low
ATG	Antigua and Barbuda	0	3	3	Very low	High	High
AUS	Australia	3	3	0	High	High	Very low
AUT	Austria	3	3	3	High	High	High
AZE	Azerbaijan	3	1	2	High	Low	Medium
BDI	Burundi	0	0	0	Very low	Very low	Very low
BEL	Belgium	3	3	3	High	High	High
BEN	Benin	1	0	0	Low	Very low	Very low
BFA	Burkina Faso	1	0	0	Low	Very low	Very low
BGD	Bangladesh	3	0	3	High	Very low	High
BGR	Bulgaria	3	2	2	High	Medium	Medium
BHR	Bahrain	2	3	3	Medium	High	High
BHS	Bahamas	1	3	2	Low	High	Medium
BIH	Bosnia and Herzegovina	1	1	1	Low	Low	Low
BLR	Belarus	3	2	1	High	Medium	Low
BLZ	Belize	0	1	0	Very low	Low	Very low
BOL	Bolivia	2	1	0	Medium	Low	Very low
BRA	Brazil	3	2	0	High	Medium	Very low
BRB	Barbados	0	3	3	Very low	High	High
BRN	Brunei	1	3	3	Low	High	High
BTN	Bhutan	3	4	3	High	Very high	High
BWA	Botswana	1	2	0	Low	Medium	Very low
CAF	Central African Republic	0	0	0	Very low	Very low	Very low
CAN	Canada	3	3	0	High	High	Very low
CHE	Switzerland	3	3	3	High	High	High
CHL	Chile	3	3	1	High	High	Low
CHN	China	4	0	3	Very high	Very low	High
CIV	Cote d'Ivoire	3	0	0	High	Very low	Very low
CMR	Cameroon	3	0	0	High	Very low	Very low
COD	Democratic Republic of Congo	3	0	0	High	Very low	Very low

COG	Congo	1	0	0	Low	Very low	Very low
COL	Colombia	3	1	1	High	Low	Low
COM	Comoros	0	0	2	Very low	Very low	Medium
CPV	Cape Verde	0	1	1	Very low	Low	Low
CRI	Costa Rica	3	2	3	High	Medium	High
CUB	Cuba	3	2	3	High	Medium	High
CYP	Cyprus	2	3	3	Medium	High	High
CZE	Czechia	3	3	3	High	High	High
DEU	Germany	3	3	3	High	High	High
DJI	Djibouti	0	1	0	Very low	Low	Very low
DMA	Dominica	0	2	2	Very low	Medium	Medium
DNK	Denmark	3	3	3	High	High	High
DOM	Dominican Republic	3	2	3	High	Medium	High
DZA	Algeria	3	1	0	High	Low	Very low
ECU	Ecuador	3	1	1	High	Low	Low
EGY	Egypt	3	1	1	High	Low	Low
ERI	Eritrea	0	0	0	Very low	Very low	Very low
ESP	Spain	3	3	3	High	High	High
EST	Estonia	2	3	2	Medium	High	Medium
ETH	Ethiopia	3	0	0	High	Very low	Very low
FIN	Finland	3	3	2	High	High	Medium
FJI	Fiji	0	1	1	Very low	Low	Low
FRA	France	3	3	3	High	High	High
FSM	Micronesia (country)	0	1	2	Very low	Low	Medium
GAB	Gabon	1	2	0	Low	Medium	Very low
GBR	United Kingdom	3	3	3	High	High	High
GEO	Georgia	1	1	1	Low	Low	Low
GHA	Ghana	3	0	1	High	Very low	Low
GIN	Guinea	1	0	0	Low	Very low	Very low
GMB	Gambia	0	0	0	Very low	Very low	Very low
GNB	Guinea-Bissau	0	0	0	Very low	Very low	Very low
GNQ	Equatorial Guinea	1	2	1	Low	Medium	Low
GRC	Greece	3	3	3	High	High	High
GRD	Grenada	0	2	3	Very low	Medium	High
GTM	Guatemala	3	1	2	High	Low	Medium
GUY	Guyana	0	2	0	Very low	Medium	Very low

HND	Honduras	2	0	1	Medium	Very low	Low
HRV	Croatia	3	3	3	High	High	High
HTI	Haiti	1	0	2	Low	Very low	Medium
HUN	Hungary	3	3	3	High	High	High
IDN	Indonesia	3	1	2	High	Low	Medium
IND	India	3	0	3	High	Very low	High
IRL	Ireland	3	3	3	High	High	High
IRN	Iran	3	0	0	High	Very low	Very low
IRQ	Iraq	3	1	1	High	Low	Low
ISL	Iceland	2	3	1	Medium	High	Low
ISR	Israel	3	3	3	High	High	High
ITA	Italy	3	3	3	High	High	High
JAM	Jamaica	1	1	3	Low	Low	High
JOR	Jordan	3	1	1	High	Low	Low
JPN	Japan	3	3	3	High	High	High
KAZ	Kazakhstan	3	2	0	High	Medium	Very low
KEN	Kenya	3	0	0	High	Very low	Very low
KGZ	Kyrgyzstan	0	0	0	Very low	Very low	Very low
KHM	Cambodia	2	0	0	Medium	Very low	Very low
KIR	Kiribati	0	0	1	Very low	Very low	Low
KNA	Saint Kitts and Nevis	0	3	3	Very low	High	High
KOR	South Korea	3	3	3	High	High	High
KWT	Kuwait	3	3	3	High	High	High
LAO	Laos	1	0	0	Low	Very low	Very low
LBN	Lebanon	1	0	3	Low	Very low	High
LBR	Liberia	0	0	0	Very low	Very low	Very low
LBY	Libya	3	1	0	High	Low	Very low
LCA	Saint Lucia	0	2	3	Very low	Medium	High
LIE	Liechtenstein	0	4	3	Very low	Very high	High
LKA	Sri Lanka	3	1	3	High	Low	High
LSO	Lesotho	0	0	0	Very low	Very low	Very low
LTU	Lithuania	3	3	3	High	High	High
LUX	Luxembourg	3	3	3	High	High	High
LVA	Latvia	2	3	2	Medium	High	Medium
MAR	Morocco	3	1	1	High	Low	Low
MDA	Moldova	1	1	1	Low	Low	Low
MDG	Madagascar	1	0	0	Low	Very low	Very low
MDV	Maldives	0	2	3	Very low	Medium	High

MEX	Mexico	3	2	2	High	Medium	Medium
MHL	Marshall Islands	0	1	3	Very low	Low	High
MKD	North Macedonia	1	1	2	Low	Low	Medium
MLI	Mali	1	0	0	Low	Very low	Very low
MLT	Malta	1	3	3	Low	High	High
MMR	Myanmar	3	0	0	High	Very low	Very low
MNE	Montenegro	0	2	1	Very low	Medium	Low
MNG	Mongolia	1	1	0	Low	Low	Very low
MOZ	Mozambique	1	0	0	Low	Very low	Very low
MRT	Mauritania	0	0	0	Very low	Very low	Very low
MUS	Mauritius	1	2	3	Low	Medium	High
MWI	Malawi	1	0	0	Low	Very low	Very low
MYS	Malaysia	3	2	3	High	Medium	High
NAM	Namibia	1	1	0	Low	Low	Very low
NER	Niger	1	0	0	Low	Very low	Very low
NGA	Nigeria	3	0	1	High	Very low	Low
NIC	Nicaragua	1	0	0	Low	Very low	Very low
NLD	Netherlands	3	3	3	High	High	High
NOR	Norway	3	3	3	High	High	High
NPL	Nepal	2	0	1	Medium	Very low	Low
NRU	Nauru	0	2	3	Very low	Medium	High
NZL	New Zealand	3	3	2	High	High	Medium
OMN	Oman	3	3	1	High	High	Low
PAK	Pakistan	3	0	1	High	Very low	Low
PAN	Panama	3	3	2	High	High	Medium
PER	Peru	3	1	0	High	Low	Very low
PHL	Philippines	3	1	3	High	Low	High
PLW	Palau	0	3	2	Very low	High	Medium
PNG	Papua New Guinea	2	0	0	Medium	Very low	Very low
POL	Poland	3	3	3	High	High	High
PRK	North Korea	1	0	0	Low	Very low	Very low
PRT	Portugal	3	3	3	High	High	High
PRY	Paraguay	2	1	0	Medium	Low	Very low
PSE	Palestine	1	1	3	Low	Low	High

QAT	Qatar	3	3	3	High	High	High
ROU	Romania	3	3	3	High	High	High
RUS	Russia	3	2	0	High	Medium	Very low
RWA	Rwanda	1	0	1	Low	Very low	Low
SAU	Saudi Arabia	3	3	1	High	High	Low
SDN	Sudan	2	0	0	Medium	Very low	Very low
SEN	Senegal	2	0	0	Medium	Very low	Very low
SGP	Singapore	3	3	4	High	High	Very high
SLB	Solomon Islands	0	0	0	Very low	Very low	Very low
SLE	Sierra Leone	0	0	0	Very low	Very low	Very low
SLV	El Salvador	2	1	3	Medium	Low	High
SOM	Somalia	0	0	0	Very low	Very low	Very low
SRB	Serbia	3	2	2	High	Medium	Medium
SSD	South Sudan	0	0	0	Very low	Very low	Very low
STP	Sao Tome and Principe	0	0	2	Very low	Very low	Medium
SUR	Suriname	0	1	0	Very low	Low	Very low
SVK	Slovakia	3	3	3	High	High	High
SVN	Slovenia	3	3	3	High	High	High
SWE	Sweden	3	3	3	High	High	High
SWZ	Eswatini	0	1	1	Very low	Low	Low
SYC	Seychelles	0	3	3	Very low	High	High
SYR	Syria	1	0	0	Low	Very low	Very low
TCD	Chad	1	0	0	Low	Very low	Very low
TGO	Togo	0	0	0	Very low	Very low	Very low
THA	Thailand	3	2	3	High	Medium	High
TJK	Tajikistan	0	0	0	Very low	Very low	Very low
TKM	Turkmenistan	3	2	0	High	Medium	Very low
TLS	Timor	0	0	0	Very low	Very low	Very low
TON	Tonga	0	1	2	Very low	Low	Medium
TTO	Trinidad and Tobago	1	3	3	Low	High	High
TUN	Tunisia	3	1	1	High	Low	Low
TUR	Turkey	3	2	3	High	Medium	High
TUV	Tuvalu	0	1	3	Very low	Low	High
TZA	Tanzania	3	0	0	High	Very low	Very low
UGA	Uganda	2	0	0	Medium	Very low	Very low
UKR	Ukraine	3	1	1	High	Low	Low
URY	Uruguay	3	3	1	High	High	Low

USA	United States	4	3	3	Very high	High	High
UZB	Uzbekistan	3	0	0	High	Very low	Very low
VCT	Saint Vincent and the Grenadines	0	2	3	Very low	Medium	High
VEN	Venezuela	3	1	0	High	Low	Very low
VNM	Vietnam	3	1	3	High	Low	High
VUT	Vanuatu	0	1	0	Very low	Low	Very low
WSM	Samoa	0	1	1	Very low	Low	Low
YEM	Yemen	1	0	0	Low	Very low	Very low
ZAF	South Africa	3	1	1	High	Low	Low
ZMB	Zambia	1	0	0	Low	Very low	Very low
ZWE	Zimbabwe	2	0	0	Medium	Very low	Very low

Source: Done by authors using Python