A Cluster Analysis on Gender Equality in the World from the Data Set of the United Nations Development Program

Hümeyra Nur Hatipoğlu

Department of Business Administration
Istanbul Commerce University, Istanbul, Turkey

Abstract-   In this study, we specifically focus on United Nations (UN) Sustainable Development Goal 5. This goal is related to gender equality. Achieving "Goal 5 Gender equality" depends on some critical variables. Also, the current situation of countries on gender equality should determine in terms of UNDP indicators. For this reason, I carried out this research to determine the current situation of 162 countries—including Turkey, regarding the principle of gender equality in the UN Sustainable Development Goal. In line with these, this study's first aim is to define the current gender inequality in the world. The second aim is the main aim for us; it is to describe Turkey's current situation in the world. The final objective is to produce practical solutions to bring closer to the "best countries group." In these aims, the researcher conducted a hierarchical clustering analysis and K-means analysis for 162 countries with 15 variables related to gender equality from United Nations Development Program (UNDP). In hierarchical clustering analysis, the researcher used Ward's method and the widely used Euclidean distance to create the distance matrix using the cluster function in R studio. Afterward, the researcher performed a k-means clustering analysis. Our findings indicate four clusters considering gender equality in terms of countries. Respectively, from best to worse, Cluster 2, Cluster 3, Cluster 4, and Cluster 1. Turkey takes place in Cluster 3. In conclusion, the researcher discussed the current situation and potential solutions for managing gender inequality.

Keywords –United Nation (UN) Sustainable Development, United Nation Development Program UNDP, Goal 5, Gender Inequality, Hierarchical Clustering, K-means Clustering Analysis.

I. INTRODUCTION

The United Nations Development Programme (UNDP) published Human Development Report considering gender equality [1]. They stated that "Human Development, if not engendered, is endangered.", so they emphasized this report's importance [2]. Although, the Human Development Report Office has yet to produce a scale that can compare countries worldwide in human development and gender inequalities [3]. However, the UN Sustainable Development Goals promote specific goals to prevent gender inequality and provide gender empowerment [4]. United Nations' sustainable development goal 5 strives to eradicate violence against women [5]. Gender inequality is a cultural, legal, and social situation that determine different rights and dignity in terms of women and man. In this situation, women and men cannot benefit from the same rights and opportunities [6]. For this reason, Goal 5's includes both realizing gender equality and empowering all women and girls [7]. Also, United Nations Development Program presents data based on gender inequality worldwide with various indicators.

The United Nations is an international organization that continues its activities in line with the UN Compact agreement with 193 member states [8]. UN has stated that its primary goal is to contribute to sustainable human development [9]. Also, there is gender equality in the ten goals (gender equality changes according to proportion to the opportunities to employees, countries' literacy rate, economic development level, etc.) [10]. Consequently, The UN efforts to solve gender inequality in 193 countries [9].

In this article, the researcher conducted a hierarchical and k-means clustering analysis on data from the UN Development Programme-Human Development Reports regarding gender equality.

In the first part of the article, the issue of gender (in)equality is addressed. Studies carried out within the scope of the UN Compact on this subject are mentioned. In addition, why gender equality is essential in social and business life and what can be done are discussed. The third section is methodology and data. In this part, the researcher introduced 15 variables handled within the scope of the UN Development Program. In the fourth section, the cluster analysis explained findings and countries categorized according to cluster analysis.   Additionally, "Elbow," "silhouette," and "fviz" methods were used to define the suitable number of clusters. The article is completed with the conclusion part.

In this paper, Ward's method and the widely used Euclidean distance in hierarchical clustering analysis it was used to create the distance matrix using the hclust function in R studio. Afterward, the researcher performed k-means clustering analysis. Our findings indicate four clusters considering gender equality in terms of countries. In this contex, my research question is "What is the situation of 162 countries on gender equality?"

Is Turkey included in which cluster? Is there a better cluster than Turkey's cluster? If so, What should Turkey and the countries in the same group do to get to the next level cluster?

II. LITERATURE REVIEW

UNDP defined some development goals for a sustainable world for all. Achieving these goals is essential for everyone to live in a sustainable world [11]. Despite UNDP's goals and work, gender inequality remains a significant problem. There have been recent developments in the area of gender inequality. However, women still make labor commitments to work more at home. They are also denied access to well-paid jobs. Therefore, they are still at a disadvantage. This gender inequality negatively affects economic development as well as the welfare of individuals [12]. Besides, in literature, gender inequality is discussed from sociological, economic, and educational perspectives. From a sociological point of view, the transformation of the biological difference between men and women into a profound and dimensional social difference threatens women's life in every sense. It forces them into subordination [13]. In terms of education, the situation of women does not draw a heartwarming picture. 2/3 of the 900 million illiterate people in the world are women. In international comparisons, it is seen that Turkey ranks 109th among 118 countries in the index, showing income inequality in terms of gender [14]. Gender inequality is seen in all areas, even in countries that implement the parliamentary system. In these countries, the number of chairs held by women is around 17% [15]. Female workers faced more unemployment and left than male workers. In particular, married female workers face much more negative outcomes than single female workers [16].

In the literature, numerous studies have been conducted on gender (in)equality, including country comparisons (e.g.[3];[12];[17]). Gender inequality brings about limitations to both the labor market and women's empowerment. For this reason, it is considered to have a negative impact on economic development [18]. Carlsen stated that to solve this problem needs to focus on education [4]. Additionally, if a government cares about Goal5, it has a better chance of re-election [17].

In research based on culture and religion, it is concluded that there is a high level of gender inequality in Islamic countries, despite fairness and justice orders [19]. Also, women in parliament and female labor force participation rates were deficient in several Islamic states[4].

In addition, the definition of gender equality has not been made much in the literature. O'Brien et al. [20] stated that gender equality has no specific definition because this term has a lot of different dimensions and indications, and these terms are pretty fractured [21]. On the other side, gender inequality defines clearly. Gender inequality offers people various opportunities due to gender-based perception [22]. In addition, this inequality is part of global inequality. According to this view, men more superior to women [23]. For this reason, gender inequality is a cultural, legal, and social situation that determine different rights and dignity in terms of women and man. In this situation, women and men cannot benefit from the same rights and opportunities [6].

As a result, gender inequality creates deep problems. Women are marginalized because of imbalance. UNDP continues its efforts to reduce discrimination against women. This research also addresses this issue precisely and offers suggestions for concrete results.

III. METODOLOGY AND DATA

A. Selection of Method

Researchers benefited from many kinds of different research methods related to gender inequality. While some researchers use qualitative research methods, some researchers utilize quantitative research methods. Koburtay et al. use qualitative methods, and they conduct content-based analysis based on religion, culture, and gender literature [19]. Also, Singh& Pandey conducted a content analysis [24]. However, Blázquez-Fernández et al. use based on panel data methods [25]. They obtain from Eurostat and the OECD statistics. In addition, they use clustering analysis on 26 countries in Europe. As a result of this, they find three clusters. Lotrič Dolinar et al. use classical clustering methods amongst the 28 European countries [26]. They use three main clusters (gender, age, and cause of

death). They analyzed by sex, age, and cause of death in terms of mortality rates in 28 European countries. However, their cluster analysis was a cluster analysis of variables, not countries. Muntele et al. use three analyses [27]; hierarchical clustering, principal component analysis, and multiple regression analysis. Also, Caous & Huarng[11] conducted a hierarchical linear model, and Conroy et al. [28] conducted a spatial Bayesian model in the USA. We conduct quantitative research methods; first, we obtain 15 variables from Human Development Data Center, parallel with gender inequality related to our research aim. Then we conduct Hierarchical Clustering. Researchers find tree clusters. After, we used K-means Clustering to four clusters in 162 countries.

Cluster analysis is used to identify subgroups or clusters in a data set. When it aggregates data in a dataset, it fulfills the purpose of separating them into different groups. This method allows close observations to be distributed in the same cluster, and dissimilar observations are distributed in different clusters[29]. In the literature, these algorithms are grouped under two headings: Hierarchical clustering techniques, Non-hierarchical clustering techniques: K-Means technique divides data sets into predetermined cluster groups. A hierarchical clustering method is a group of stages that show a tree-like structure by deleting or adding an element from the clusters. When the general purpose of cluster analysis is considered, it is explained as bringing together similar data and providing benefits to the researcher [30]. Also, in k-means analysis, we determine the number of clusters, but in hierarchical clusters, we cannot predict the number of clusters. Therefore, we draw a tree-like dendrogram for clusters from 1 to n[29]. Here, we first performed a hierarchical clustering analysis and found the number of clusters. Dendrogram results showed us that 3 clusters are more suitable. To consolidate and control our results, we performed the k-means clustering analysis. However, the results of the K-means and hierarchical clustering analyses did not come out in parallel. K-means cluster analysis showed that four clusters are more suitable for this research.

B. Introduce Data and Variables
Human Development Data Center is a platform with many datasets. This platform can access various data regarding year, country, and gender variables. The data contained here is obtained from international data organizations that have the authority, resources, and expertise to collect national data on specific indicators. This site also offers more than 150 global indicators for over 190 countries [31].
15 variables are represented below (some for both female and male):

Human Development Index (HDI) (2019): Human Development Index (HDI) includes life expectancy, education, and gender development. HDI gauges the achievement of a country in fundamental human capabilities on average [2]. For this reason, it is a composite index [31]. Also, The research on 85 countries shows that a country with a high HDI rate predicts less national corruption in this country[32].
Coefficient of human inequality (2019): Three fundamental indicators of HD (human development) are health, education, and income. The coefficient of human equality includes the average of this inequality of three indicators [31].
Inequality in education (%) (2019): This variable's data was obtained from household surveys and mentioned inequality in the distribution of years of schooling. Education inequality affects income inequality. Also, gender inequality in society decreases per capita income and increases education inequality between boys and girls [33].
Inequality in income (%) (2019): This variable's data was obtained from household surveys and mentioned inequality in income distribution [31].
Inequality in life expectancy (%) (2019): This variables' data obtain from life tables estimated about inequality in distribution of expected length of life[34]. Studies show that If people have low socioeconomic status, they have low life expectancy [35].
Life expectancy at birth (female and male (years)) (2019): These two variables (male/female)are mentioned number of years a newborn baby anticipated to live prevailing model. This model based on age mortality percentage. Also, at the time of birth remains the same situation during the life of an infant [36].
A cross-country study found that in developed countries such as Europe, if income inequality occurs, it does not significantly adversely affect health. However, this result was found for developed countries. Therefore, health policies should be given importance [25]. In a study conducted in Hungary, inequality in life expectancy is increasing. Health behaviors and inequalities in people's access to care are all closely related to inequalities in life expectancy[37]. Besides, long life expectancy of the newborn infant (male/female) is highly correlated with high per capita income, the education of the society, public expenditures for environmental and social protection [38] and quality of the public health infrastructure [27]. Fujii (2018) stated that private health expenditures have a higher impact than public health expenditures [39].

Gender Inequality Index (GII) (2019): It is a composite indicator that reflects the inequality of achievement between women and men in three dimensions. These dimensions are reproductive health, empowering women and girls, and women's place in the labor market [31]. This index shows gender-based disadvantages in the light of data. These disadvantages are reproductive health, empowerment and the labor market. The index shows the potential difference between the gains achieved by men and women in these three dimensions of disadvantage [31]. According to the index, reproductive health indicators are maternal mortality rate, adolescent birth rate; women and girls empowerment indicators, percentage of women's seats in parliament, percentage of the population with at least secondary education; The labor market indicator is labor force participation by gender. A low index means that inequality between women and men is low [31].

Income index (2019): This index is obtained using GNP per capita (2017 PPP International $, using natural logarithms), a minimum of $100 and a maximum of $75,000 [40].

Studies have found some factors affecting GNP per capita. According to Islam (2020), as the literacy rate increases, so does the GNP per capita[41].

Estimated gross national income per capita female and male (2017 PPP $) (2019): This index is derived from the ratio of women's wages to men's wages. That is, this index is produced from the shares of women and men in the economically working population and GNP [31].

Mandatory paid maternity leave (days) (2019): This index represents the minimum number of days required to be paid by both the government and the institution. It also mentions the mandatory leave that only women take according to this index; does not include men [31]. According to Broadway et al(2020), mothers who are on leave in the first year have a higher rate of returning to work [42]. However, when mandatory leave is given to men, it is stated that men intend to invest in human capital such as seeking a new job and self-development [43].

Share of employment in non-agriculture (female (% of total employment in nonagriculture)) (2019): It refers to the share of women working in industry and service fields other than agriculture in employment [31]. Looking at the classical economic development literature; if non-farm employment is increasing in an economy; Economic growth occurs when employment in agriculture decreases [44].

Share of seats in parliament (% held by women) (2019): The proportion of seats held by women in the national parliament is also a percentage of the total number of seats in the parliament [45]. According to a study conducted in the Finnish and Danish parliaments in 2019; stated that both party groups and informal organizations are working to increase women's seats in the parliament [46]. In the same year, it was found that genuine savings increased with the participation of women in the parliament. If the proportion of women in a parliament increases by 10%, there will be an increase of approximately 1.25 percentage points in the genuine savings rate [47]. Altuzarra et al. [48], on the other hand, stated that the working rate of women is generally lower. In addition, in their country-based comparison, they found a significant relationship between women's participation in parliament and economic growth in developing countries. However, this is not valid for bilateral sub-Saharan countries. In sub-Saharan countries, women participate in political life, but they do not have an impact on their political priorities and economic growth, and they face obstacles. In African countries, it has been observed that the rate of women taking part in the parliament is quite high [49].

Total unemployment rate (female to male ratio) (2019): This variable represents the percentage of women over the age of 15 who are ready to work and looking for a job, compared to men over the age of 15 with the same characteristics [50]. Although gender equality is a fundamental right, the proportion of women in the workforce is still low. There is gender inequality in business life in all countries, but this rate varies on a country-by-country basis [51]. Unfortunately, during the coronavirus epidemic, the unemployment rate for women was 24% higher than for men. In addition, women expected their salaries to be 50% lower than men [52]. In addition, some specific reasons were found for the higher unemployment rates of women. These are: university education rates and housing costs[28]. For these reasons, women had to face higher rates of unemployment [16].

IV. FINDINGS

A. Hierarcial Cluster Analysis

Hierarchical clustering does not require us to stick to a particular choice of K. In addition, clusters are revealed thanks to the dendrogram[29]. In Table 1, we calculated the Coefficient of variation to define the compactness of the data.

Table 1: Selection of Distance

| > sapply(country, function(x) sd(x) / mean(x))	CV
Coefficient of human inequality (2019)	0.6617657
Estimated gross national income per capita female (2017 PPP $) (2019)	1.0725343
Estimated gross national income per capita male (2017 PPP $) (2019)	0.9849815
Gender Inequality Index (GII)(2019)	0.5520799
Human Development Index (HDI) (2019)	0.2137279
Income index (2019)	0.2482683
Inequality in education (%) (2019)	0.7955322
Inequality in income (%) (2019)	0.6660143
Inequality in life expectancy (%) (2019)	0.8640804
Life expectancy at birth (female (years)) (2019)	0.4072212
Life expectancy at birth (male (years)) (2019)	0.3939358
Mandatory paid maternity leave (days) (2019)	0.5587116
Share of employment in nonagriculture (female (% of total employment in nonagriculture)) (2019)	0.2903126
Share of seats in parliament (% held by women) (2019)	0.4911014
Total unemployment rate (female to male ratio) (2019)	1.0059038

The majority of CV values are below 1. So the data is relatively compact. That is why we used the squared euclidean distance. Also, cluster analysis aims to create homogeneous clusters as much as possible. Ward's Method will determine the main clusters best. Ward's Method was preferred in this article because it was decided to investigate the main groups that provide division without aiming to obtain small groups.

Applying Hierarchical Clustering

Cluster method:ward.D
Distance: squared euclidean
Number of objects: 162

Cluster Dendrogram

Figure 1: Cluster Dendogram
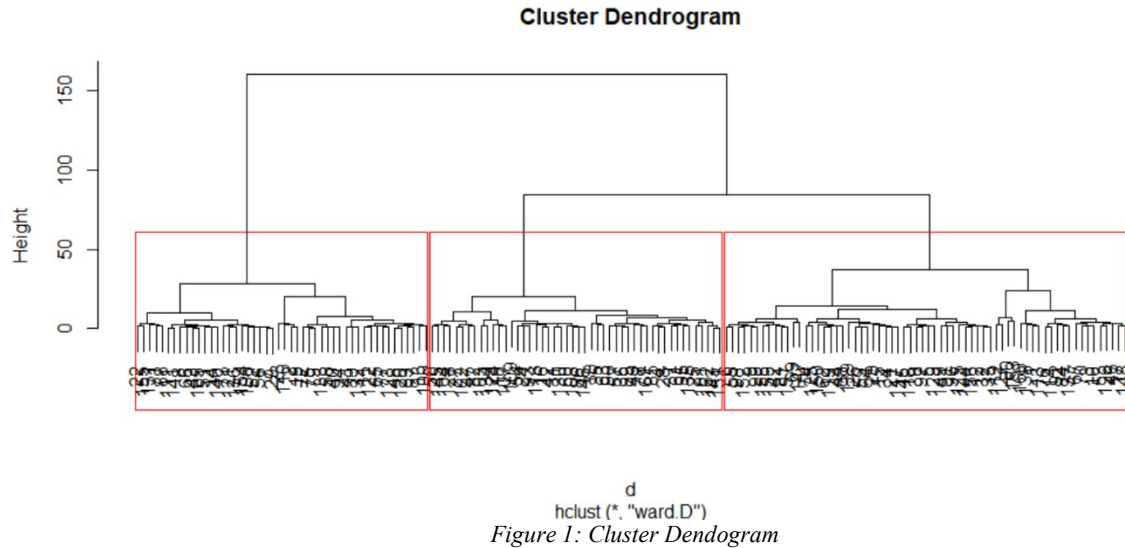
The researcher derived data from 162 countries and 15 variables and divided clusters for each observation in red. Figure 1: Cluster Dendrogram shows that each leaf of the dendrogram represents one of 162 observations.

As they climb up the tree, the leaves turn into branches. Similar observations come together. The higher it goes, the more the branches come together. That is, observations that are similar to each other continue to accumulate. The observations that converge at the bottom of the tree are similar. However, the observations that converge near the top of the tree are pretty different [29]. According to Figure 1, it is seen that 3 clusters are suitable. This figure can be seen more clearly in Figure 2: Clusplot.
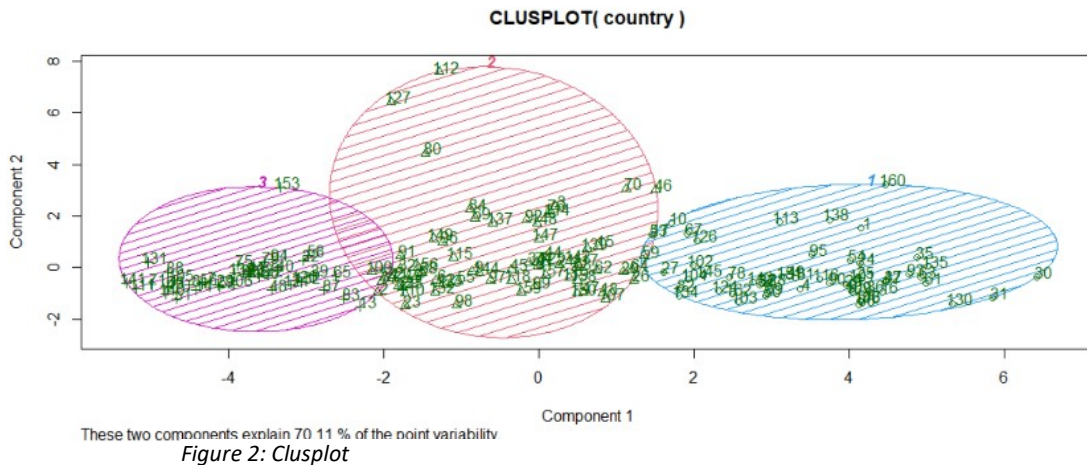
,

Figure 2: Clusplot

Figure 2: According to the clusplot, 3 clusters are clearly seen. However, the cluster with the red color coincided with the 0 points. For this reason, it is recommended to try 2 or 4 clusters. This recommendation was taken into account at the k-means stage. The characteristics of the 3 clusters will be examined at the moment according to the average cluster values in Table 2.

Tablo 2: Hierarchical Clustering Table (for 3 clusters)

Clusterk 3 (Hierarchical Cluster)	1		2		3	
Coefficient of human ine~	18,53	Medium	8,83	Low	30,06	High
Estimated gross national income(f)	9109	Medium	33964	High	2913	Low
Estimated gross national income (m)	19380	Medium	51042	High	4359	Low
Gender Inequality Inde~	0,38	Medium	0,12	Low	0,55	High
Human Development Inde~	0,73	Medium	0,89	High	0,54	Low
Income index (20~	0,72	Medium	0,90	High	0,51	Low
Inequality in educatio~	17,17	Medium	6,39	Low	34,07	High
Inequality in income	24,94	Medium	16,11	Low	28,48	High
Inequality in life expec~	12,58	Medium	4,33	Low	27,57	High
Life expectancy at birth (female)	76,46	Medium	82,62	High	65,26	Low
Life expectancy at birth (male)	71,35	Medium	77,26	High	61,26	Low
Mandatory paid maternity~	104,80	Medium	138,44	High	90,08	Low
Share of employment in nonagriculture (female)	35,20	Low	46,17	High	45,43	Medium
Share of seats in parliament (female)	20,62	Low	28,49	High	23,05	Medium
Total unemployment rate (female)	1,83	High	1,05	Low	1,14	Medium

*Cluster 1:*Countries included in Cluster 1: Coefficient of human inequality, Estimated gross national income per capita female/male, gender inequality, Human Development Index, Income index, Inequality in education, Inequality in life expectancy, Inequality in income index, Life expectancy at birth (female/male (years)), Mandatory paid maternity leave (days) are medium; Share of employment in non-agriculture (female (% of total employment in non-agriculture) and Share of seats in parliament(women) are low;   Total unemployment rate (female to male ratio) is high.

These countries are: Afghanistan, Algeria, Argentina, Armenia, Azerbaijan, Bangladesh, Barbados, Belize, Bhutan, Botswana, Brazil, Cabo Verde, Chile, China, Colombia, Costa Rica, Dominican Republic, Ecuador, Egypt, Gabon, Georgia, Guatemala, Guyana, Honduras, India Indonesia, Iran (Islamic Republic of), Iraq, Jamaica, Jordan, Kuwait, Kyrgyzstan, Lebanon, Malaysia, Maldive,s Mauritius, Mexico, Moldova (Republic of), Mongolia, Oman, Pakistan, Palestine(State of), Panama, Paraguay, Peru, Philippines, Samoa, Sao Tome and Principe, Saudi Arabia, Serbia, Sri Lanka, Sudan, Suriname, Tajikistan, Thailand, Tonga, Tunisia, Turkey, Turkmenistan, Ukraine, United Arab Emirates, Uruguay, Uzbekistan, Venezuela (Bolivarian Republic of), Viet Nam, and Yemen.

Cluster 2:   Countries included in Cluster 2: Coefficient of human inequality, gender inequality, Inequality in education, Inequality in life expectancy, Inequality in income index, and Total unemployment rate (female to male ratio) are low; Estimated gross national income per capita female/male, Human Development Index, Income index, Life expectancy at birth (female/male (years)), Mandatory paid maternity leave (days), hare of employment in non-agriculture (female (% of total employment in non-agriculture) and Share of seats in parliament(women) are high.

These countries are Albania, Australia, Austria, Bahamas, Belarus, Belgium, Bosnia and Herzegovina, Brunei, Darussalam, Bulgaria, Canada, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany Greece Hungary Iceland Ireland Israel Italy Japan Kazakhstan, Korea (Republic of), Latvia, Lithuania, Luxembourg, Malta, Montenegro, Netherlands, New Zealand, North Macedonia, Norway, Poland, Portugal, Romania, Russian Federation, Singapore, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom and United States.

__Cluster 3__: Countries included in Cluster 3: Coefficient of human inequality, gender inequality, Inequality in education, and inequality in life expectancy are medium; Inequality in income index, Share of employment in non-agriculture (female (% of total employment in non-agriculture) are medium; Total unemployment rate (female to male ratio) is very high; lastly, Estimated gross national income per capita female/male, Human Development Index, Income index, Life expectancy at birth (female/male (years)), Mandatory paid maternity leave (days), and Share of seats in parliament(women) are high.

These countries are Angola, Benin, Bolivia (Plurinational State of), Burkina Faso, Burundi, Cambodia, Cameroon, Central African Republic, Chad, Comoros, Congo, Congo (Democratic Republic of the), Côte d'Ivoire, El Salvador, Eswatini (Kingdom of), Ethiopia, Gambia, Ghana, Guinea, Guinea-Bissau, Haiti, Kenya, Lao People's Democratic Republic, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mozambique, Myanmar, Namibia, Nepal, Nicaragua, Niger, Nigeria, Papua, New Guinea, Rwanda, Senegal, Sierra Leone, South Africa, South Sudan, Tanzania (United Republic of), Timor-Leste, Togo, Uganda, Zambia and Zimbabwe.

B. K-means Cluster Analysis

K-means clustering is an approach that divides a dataset into K-disparate and non-overlapping clusters. The number of K clusters must be determined to use this method. This is also explained as a potential disadvantage [29].Thus, it takes many trials until it finds the right cluster. However, since we had hierarchical clusters before, we could determine the appropriate number of clusters with only three attempts by doing K-1 and K+1.

two properties of clusters:

-Each of the 162 observations belongs to at least one cluster ($C_1 \cup C_2 \cup \ldots \cup C_{15} = \{1, \ldots, 162\}$)

- No favorites belong to more than one cluster ($C_{15} \cap C_{15'} = \emptyset$ for all $k_{15'} \neq k_{15'}$)

According to the K-means clustering method, 3 clusters were applied based on the theory. Here, the countries were expected to be divided into three: developed, developing, and underdeveloped. However, as seen in the table below, 3 clusters were unsuitable for our data set. Examining the plots for all three cases:
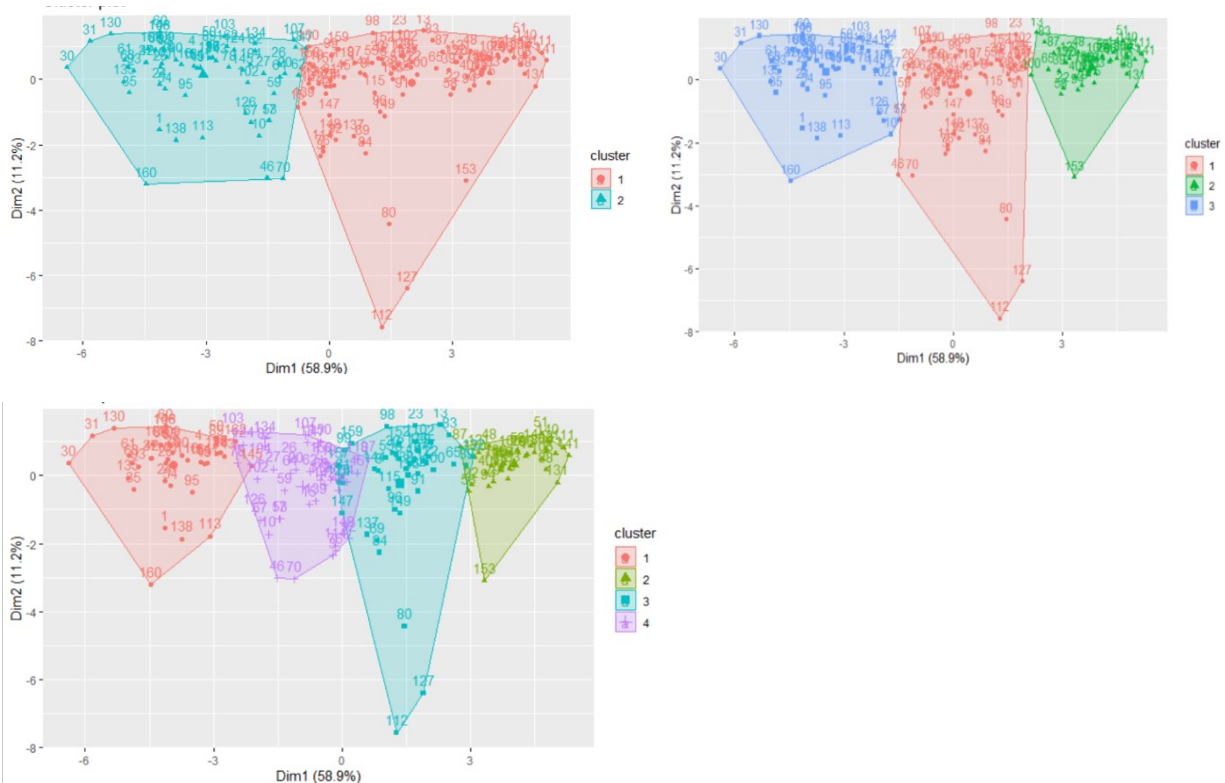
The color of each observation represents the cluster to which it has been assigned using the K-means clustering algorithm. There is no ordering of the clusters [29]. In Figure 2, Figure 3, and Figure 4, we see a simulated data set with 162 observations in two-dimensional space. When Figures 1, 2, and 3 are examined, it is seen that the best separation is made with 2 and 4 clusters. For this reason, the optimum profit was sought using the Elbow and Silhouette Methods, respectively. The results of the Elbow and the Silhouette method are presented in Cluster 2 and Cluster 4, unlike hierarchical clustering.

Determination of optimum k with Elbow and Silhouette Method

The elbow method is one of the most widely used methods to discriminate the optimal number of clusters. The oldest visual method for estimating the optimal number of clusters for the analyzed dataset, the Elbow method, often requires K-means to occur on the same dataset with a contiguous range of cluster numbers. Experienced analysts predict the optimum elbow point by analyzing the elbow curve. This point represents the potential optimal number of clusters [53]. The Silhouette Method is a method of interpreting and validating data contained in a cluster. This method makes a graphical representation of how well each observation is found in its cluster[54].
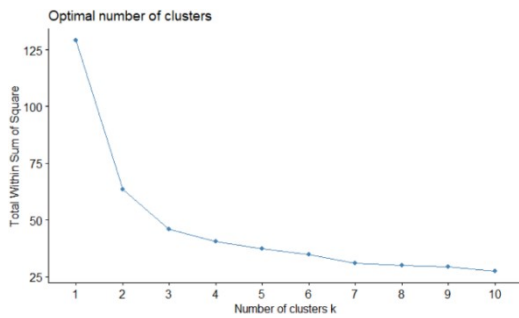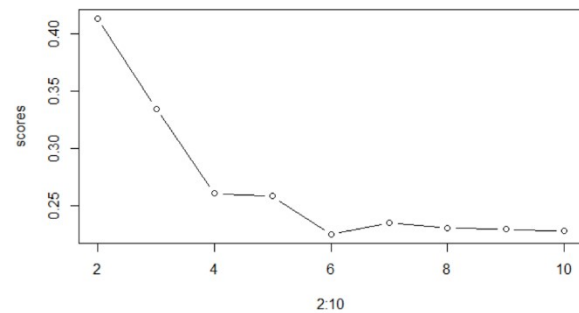

Figure 6: Elbow Method

Figure 7: Silhoutte Method

We see that k is not 3 but 2 or 4 in both elbow and Silhouette methods. Let's try with Fviz for a more precise distinction. Figure 6 contains the Fviz results.
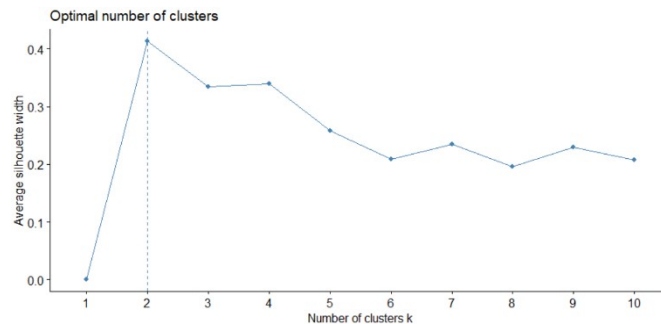

Figure 8: Fviz Method

According to the Fviz result, the highest value appears when 2 clusters are. Secondly, it appears when the highest value is 4 clusters. We take k=4 to make a more precise clustering. When we perform the k-means analysis again for K=4, we see the following clustering. The properties for 4 clusters are presented in Table 3.

Table 3:K-Means Clustring

Cluster k 4 (K-means)	1		2		3		4	
Coefficient of human ine~	31.6	very high	8.71	low	12.5	medium	22.3	high
Estimated gross national income(f)	2140.	low	39683.	very high	14255.	high	6367.	medium
Estimated gross national income (m)	3545.	low	60435.	very high	25973.	high	12739.	medium
Gender Inequality Inde~	0.582	very high	0.0886	low	0.270	medium	0.437	high
Human Development Inde~	0.510	low	0.917	very high	0.794	high	0.689	medium
`Income index (20~	0.483	low	0.934	very high	0.783	high	0.672	medium
`Inequality in educatio~	37.1	very high	5.88	low	9.02	medium	22.7	high
`Inequality in income	27.7	high	16.2	low	20.4	medium	28.2	very high
`Inequality in life expec~	29.8	very high	3.75	low	7.93	medium	15.7	high
`Life expectancy at birth (female)	63.8	low	83.9	very high	78.6	high	74.6	medium
`Life expectancy at birth (male)	60.3	low	79.2	very high	72.7	high	69.5	medium
`Mandatory paid maternity~	88.7	low	111.	high	141.	very high	98.9	medium
Share of employment in nonagriculture (female)	42.2	high	45.5	very high	41.5	medium	37.4	low
Share of seats in parliament (female)	19.6	low	30.7	very high	20.9	medium	24.7	high
Total unemployment rate (female)	1.22	medium	1.15	low	1.68	very high	1.48	high

Cluster 1:Countries included in Cluster 1: Coefficient of human inequality, gender inequality, Inequality in education, inequality in life expectancy is very high; Inequality in income index, Share of employment in non-agriculture (female (% of total employment in non-agriculture) are high; Total unemployment rate (female to male ratio) is medium; lastly, Estimated gross national income per capita female/male, Human Development Index, Income index, Life expectancy at birth (female/male (years)), Mandatory paid maternity leave (days), and Share of seats in parliament(women) are low.

These countries are Afghanistan, Angola, Benin, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Comoros, Congo, Congo (Democratic Republic of the), Côte d'Ivoire, Eswatini (Kingdom of), Ethiopia, Gambia, Ghana, Guinea, Guinea-Bissau, Haiti, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mozambique, Niger, Nigeria, Pakistan, Papua New Guinea, Senegal, Sierra Leone, South Sudan, Sudan, Tanzania (United Republic of), Timor-Leste, Togo, Uganda, Yemen, Zambia and Zimbabwe.

Cluster 2: Countries included in Cluster 2: Coefficient of human inequality, gender inequality, Inequality in education, inequality in life expectancy is low; Inequality in income index, Share of employment in non-agriculture (female (% of total employment in non-agriculture) are low; Total unemployment rate (female to male ratio) is low; lastly, Estimated gross national income per capita female/male, Human Development Index, Income index, Life expectancy at birth (female/male (years)), Mandatory paid maternity leave (days) and the Share of seats in parliament(women) are high/very high.

These countries are Australia, Austria, Belgium, Brunei Darussalam, Canada, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Japan, Korea (Republic of), Lithuania, Luxembourg, Malta, Netherlands, New Zealand, Norway, Poland, Portugal, Singapore, Slovenia, Spain, Sweden, Switzerland, United Arab Emirates, United Kingdom and United States.

Cluster 3:Countries included in Cluster 3: Coefficient of human inequality, gender inequality, Inequality in education, and inequality in life expectancy are medium; Inequality in income index, Share of employment in non-agriculture (female (% of total employment in non-agriculture) are medium; Total unemployment rate (female to male ratio) is very high; lastly, Estimated gross national income per capita female/male, Human Development Index, Income index, Life expectancy at birth (female/male (years)),   Mandatory paid maternity leave (days), and Share of seats in parliament(women) are high.

These countries are Albania, Argentina, Armenia, Azerbaijan, Bahamas, Barbados, Belarus, Bosnia and Herzegovina, Bulgaria, Chile, China, Costa Rica, Croatia, Georgia, Hungary, Iran (Islamic Republic of), Jamaica, Kazakhstan, Kuwait, Kyrgyzstan, Latvia, Lebanon, Malaysia, Mauritius, Moldova (Republic of), Mongolia, Montenegro, North Macedonia, Oman, Panama, Romania, Russian Federation, Samoa, Saudi Arabia, Serbia, Slovakia, Sri Lanka, Thailand, Tonga, Turkey, Ukraine, United Kingdom, Uzbekistan and Viet Nam.

Cluster 4:Countries included in Cluster 4: Countries that include Cluster 4 are Coefficient of human inequality, gender inequality, Inequality in education, inequality in life expectancy high; Inequality in income index, Share of employment in non-agriculture (female (% of total employment in non-agriculture) are high; Total unemployment rate (female to male ratio) is high; lastly, Estimated gross national income per capita female/male, Human Development Index, Income index, Life expectancy at birth (female/male (years)),   Mandatory paid maternity leave (days), and Share of seats in parliament(women) are medium.

These countries are Algeria, Bangladesh, Belize, Bhutan, Bolivia (Plurinational State of), Botswana, Brazil, Cabo Verde, Cambodia, Colombia, Dominican Republic, Ecuador, Egypt, El Salvador, Gabon, Guatemala, Guyana, Honduras, India, Indonesia, Iraq, Jordan, Kenya, Lao People's Democratic Republic, Maldives, Mexico, Myanmar, Namibia, Nepal, Nicaragua, Palestine State of, Paraguay, Peru, Philippines, Rwanda, Sao Tome and Principe, South Africa, Suriname, Tajikistan, Tunisia, Turkmenistan and Venezuela (Bolivarian Republic of).

V. Conclusion and Recommendation

In this study, we specifically focus on United Nations (UN) Sustainable Development Goal 5. This goal is related to gender equality. Achieving "Goal 5 Gender equality" depends on some critical variables. The UN Sustainable Development Goals promote specific goals to prevent gender inequality and provide gender empowerment [4]. United Nations sustainable development goal 5 strives to eradicate violence against women[5]. Goal 5's includes both realizing gender equality and empowering all women and girls [7]. Also, gender inequality is part of global inequality. According to this view, men more superior to women [23]. The current research aims to see the current status of this gender inequality in 162 countries. For this reason, this research was carried out to determine the current state of 162 countries, especially Turkey, regarding the principle of gender equality in the UN SDG. Following this, this study's first aim is to define the current gender inequality in the world. The second aim is mainly for us; it is to define Turkey's current situation in the world. The final aim is to produce practical solutions to bring closer to the "best countries group."

The clustering method is widely used and is divided into hierarchical and non-hierarchical clustering. In the K-means clustering (non-hierarchical) method, clusters must be determined beforehand. For this reason, we first applied hierarchical clustering analysis. According to the hierarchical cluster analysis, 162 countries were divided into three categories regarding gender equality. The countries with the best features in these three categories attach importance to women's rights and gender equality. Afterward, a K-means analysis was performed. In the K-means analysis, we used one more and one less cluster than the cluster we found in the hierarchical cluster analysis. According to the results of the K-means analysis, we found the optimum number of clusters for gender equality to be four. Looking at the results of the research, it is possible to categorize clusters 2, 3, 4, and 1 in terms of gender equality, respectively. Cluster 2 is also the country with the highest income levels and the lowest expectations for inequality in education, income, and life expectancy. In Cluster 2, the number of female members of the Parliament is quite high compared to other clusters, and the female unemployment rate is low. The cluster with the highest

inequality is cluster 1. It is seen that the results in this cluster are diametrically opposite to Cluster 2. In Cluster 1, both men and women have low-income levels, very high gender inequality, and high expected inequality in education, income, and life. In addition, there is low participation of women in management in cluster 3 (Turkey include too). When Cluster 3 is examined, it is seen that men's income is higher than women's, women's unemployment rate is very high, and the expectation of inequality in education, income, and living level is medium. In this respect, it is in a worse position than cluster 4, but it can be said that it is in a better position than cluster 4 in terms of income index. The representation of women in Parliament is higher in Cluster 4 than in Cluster 3. Cluster 3 and Cluster 4 exhibit similar features in many features. However, the important thing is to be in something other than clusters 3 or 4. Whether in Cluster 3 or Cluster 4, the country's goal should be to be in Cluster 2. Turkey, which is in Cluster 3, should also make arrangements in business, family life, and social life to be included in Cluster 2 immediately. Countries included in cluster 2 are Australia, Austria, Belgium, Brunei Darussalam, Canada, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Japan, Korea (Republic of), Lithuania, Luxembourg, Malta, Netherlands, New Zealand, Norway, Poland, Portugal, Singapore, Slovenia, Spain, Sweden, Switzerland, United Arab Emirates, United Kingdom, and United States. When four clusters were preferred, Turkey was also included in the cluster of countries that attach importance to gender equality at an average level.

Looking at the countries with the best level of gender equality in both analyzes, when four clusters are determined, it is seen that the United Arab Emirates is among the countries that give importance to gender equality. On the other hand, in research based on culture and religion, it is concluded that there is a high level of gender inequality in Islamic countries, despite fairness and justice orders [19]. Also, women in parliament and female labor force participation rates were deficient in a number of Islamic states [4]. On the other hand, the fact that the rate of women in parliament is 50% is supported by the data. However, it is seen that only 35 countries out of 162 countries give due importance to gender equality. Unfortunately, the remaining 12 countries still have many steps to take. Turkey is among 127 countries. According to HDR results, it is ranked 54th in the Gender Development Index [31].

The results of the research suggest some methods. Such as, Infrastructure factors such as transportation routes affect non-agricultural employment. Infrastructure work has been observed to have a strong positive effect on women's non-agricultural employment in societies with developed transport infrastructure and accepting more egalitarian rules [55]. This has shown us that applying a few methods is not enough to overcome gender inequality. Equality will be possible with multiple methods and development in all areas. In addition, In a study on fertility rates and women's employment, child-rearing came to the fore as motherhood was seen as the primary duty of women, even if the fertility rate decreased. Traditional gender norms have prioritized women's investment in their children's overabundant opportunities outside the home [56]. For this reason, it is necessary to divide the responsibilities of the child as parents. The income per capita increases positively with the increase in the adult literacy rate. Thanks to the investments made in education, the adult literacy rate increased. A society's higher literacy rate results in better jobs, higher earnings, and healthier behavior. This supports healthy life expectancy. This support increases productivity. For all these reasons, more expenditure should be made on education [41]. States have a great responsibility here. The state should provide free education in poor areas and employ employment programs to reduce income and education inequality, respectively. States need to build more schools, especially for girls, to eliminate gender inequality in education. In addition, it should try to reduce gender inequality by starting online education programs today [33]. At the same time, it is necessary to encourage young people to participate in entrepreneurship programs to prevent unemployment [57]. Focusing on public programs that are more relevant to the population's health status, such as spending on both environmental and social protection, will have more impact [38]. Gender quotas should be enforced in parliament. This way, significant government representation is ensured [49]. Despite all the difficulties, for example, companies in India are increasing in women-friendly policy practices [24]. This way, company cases should be examined and taken as an example.

Research Limitations
This research is based on 2019 data. Therefore, the Covid-19 effect is not seen in the study. Consequently, it is recommended to repeat. In addition, the data of this study were obtained from UNDP. A more comprehensive analysis can be done by using the reports published by companies, especially for women working in the business world.

Practical Implications

Our findings clearly outline the features necessary for a better, just, and equitable world. Because with gender equality, economic growth will be achieved, and the way to achieve sustainable development goals will be opened. In addition, with this study, countries can see where they are. The current situation should be determined first if you want to follow a path. This study, intended to determine the current situation, sheds light on 162 countries.

REFERENCES

[1] I. Permanyer, "A Critical Assessment of the UNDP's Gender Inequality Index," Feminist Economics, vol. 19, (2), pp. 1-32, 2013.

[2] (UNDP, 1995, p.1)

[3] S. Klasen, "UNDP's gender-related measures: Current problems and proposals for fixing them," 2017.

[4] L. Carlsen, "Gender inequality and development," Sustainability Science, vol. 15, (3), pp. 759-780, 2020.

[5] J. Sebire, "Why gender equality in policing is Important for achieving United Nations sustainable development goals 5 and 16," International Journal for Crime, vol. 9, (1), pp. 80-85, 2020.

[6] European Institute for Gender Equality Resport, 2020

[7]The Sustainable Development Goals Report, 2020

[8] United Nations, https://www.un.org/en/about-us/

[9] UNDP Annual Report, 2008

[10] UN Women, Gender equality and sustainable development, 2014

[11] E. L. Caous and F. Huarng, "Economic complexity and the mediating effects of income inequality: Reaching sustainable development in developing countries," Sustainability, vol. 12, (5), pp. 2089, 2020.

[12]S. Dilli, S. G. Carmichael and A. Rijpma, "Introducing the historical gender equality index," Feminist Economics, vol. 25, (1), pp. 31-57, 2019.

[13] U. Eliuz, "Panic of sexual idenity: Being a woman abstract cinsel kimlik paniği: Kadin olmak" Turkish Studies, 2011.

[14] World Economic Forum (2007). The Global Gender Gap Report http://www3.weforum.org/docs/WEF_GenderGap_Report_2007.pdf (date of Access: 25 May 2022)

[15] TÜSİAD, Faaliyet Raporu, 2008

[16] S. Ham, "Explaining Gender Gaps in the South Korean Labor Market During the COVID-19 Pandemic," Feminist Economics, vol. 27, (1-2), pp. 133-151, 2021.

[17] C. Kroll and V. Zipperer, "The political economy of the sustainable development goals (SDGs): Does performance on the SDGs affect re-election?" Sustainability (Basel, Switzerland), vol. 12, (16), pp. 1-14, 2020.

[18] T. M. H. Bui, X. V. Vo and D. T. Bui, "Gender inequality and FDI: empirical evidence from developing Asia–Pacific countries," Eurasian Economic Review, vol. 8, (3), pp. 393-416, 2018.

[19] T. Koburtay, J. Syed and R. Haloub, "Implications of Religion, Culture, and Legislation for Gender Equality at Work: Qualitative Insights from Jordan," Journal of Business Ethics, vol. 164, (3), pp. 421-436, 2020.

[20] K. R. O'Brien et al, "What Is Gender Equality in Science?" Trends in Ecology & Evolution (Amsterdam), vol. 34, (5), pp. 395-399, 2019.

[21] M. Barnett, "Gender equality, norms and practices: Post-script to special issue on new actors, old donors and gender equality norms in international development cooperation," Progress in Development Studies, vol. 18, (3), pp. 208-213, 2018.

[22] R. W. Kolb, The SAGE Encyclopedia of Business Ethics and Society. 20187. DOI: 10.4135/9781483381503.

[23] F. Tomos et al, Women Entrepreneurs and Strategic Decision Making in the Global Economy. 2019.

[24] S. Singh and M. Pandey, "Women-friendly policies disclosure by companies in India," Equality, Diversity and Inclusion, vol. 38, (8), pp. 857-869, 2019.

[25] Blázquez-Fernández, C., Cantarero-Prieto, D. & Pascual-Saez, M. 2018, "Does Rising Income Inequality Reduce Life Expectancy? New Evidence for 26 European Countries (1995-2014)", Global economic review, vol. 47, no. 4, pp. 464-479.

[26] A. Lotrič Dolinar, J. Sambt and S. Korenjak-Černe, "Clustering EU Countries by Causes of Death," Population Research and Policy Review, vol. 38, (2), pp. 157-172, 2019.

[27] I. Muntele et al, "Trends in life expectancy in Romania between 1990 and 2018. A territorial analysis of its determinants," Sustainability, vol. 12, (9), pp. 3802, 2020.

[28] T. Conroy, S. Deller and P. Watson, "Regional income inequality: a link to women-owned businesses," Small Business Economics, vol. 56, (1), pp. 189-207, 2021;2019;

[29] James, G. et al "An introduction to statistical learning" Vol. 112, 2013 New York: springer.

[30] D. J. Ketchen and C. L. Shook, " The Application of Cluster Analysis in Strategic Management Research: An Analysis and Critique," Strategic Management Journal, vol. 17, (6), pp. 441-458, 1996.

[31] HDR, 2020

[32] W. Scholl and C. C. Schermuly, "The Impact of Culture on Corruption, Gross Domestic Product, and Human Development," Journal of Business Ethics, vol. 162, (1), pp. 171-189, 2020.

[33] K. Munir and A. Kanwal, "Impact of educational and gender inequality on income and income inequality in South Asian countries," International Journal of Social Economics, vol. 47, (8), pp. 1043-1062, 2020.

[34] wpp, 2018

[35] G. T. Sanzenbacher et al, "Rising Inequality in Life Expectancy by Socioeconomic Status," North American Actuarial Journal, vol. 25, (S1), pp. S566-S581, 2021.

[36] wpp, 2019

[37] A. Bíró et al, "Life expectancy inequalities in Hungary over 25 years: The role of avoidable deaths," Population Studies, vol. 75, (3), pp. 443-455, 2021.

[38] P. A. M. Cervantes, N. R. López and S. C. Rambaud, "The relative importance of globalization and public expenditure on life expectancy in europe: An approach based on mars methodology," International Journal of Environmental Research and Public Health, vol. 17, (22), pp. 1-20, 2020.

[39] T. Fujii, "Sources of health financing and health outcomes: A panel data analysis," Health Economics, vol. 27, (12), pp. 1996-2015, 2018.

[40] World Development Indicators database

[41] M. S. Islam, "Human Capital and Per Capita Income Linkage in South Asia: A Heterogeneous Dynamic Panel Analysis," Journal of the Knowledge Economy, vol. 11, (4), pp. 1614-1629, 2020.

[42] B. Broadway et al, "The Impact of Paid Parental Leave on Labor Supply and Employment Outcomes in Australia," Feminist Economics, vol. 26, (3), pp. 30-65, 2020.

[43] D. T. Tharp and E. J. Parks-Stamm, "Gender Differences in the Intended Use of Parental Leave: Implications for Human Capital Development," Journal of Family and Economic Issues, vol. 42, (1), pp. 47-60, 2021.

[44] J. D. Merfeld, "Moving Up or Just Surviving? Nonfarm Self-Employment in India," American Journal of Agricultural Economics, vol. 102, (1), pp. 32-53, 2020.

[45] IPU Parline, https://data.ipu.org/women-ranking (date of Access: 1 May 2022)

[46] J. Kantola and L. Rolandsen Agustín, "Gendering the Representative Work of the European Parliament: A Political Analysis of Women MEP's Perceptions of Gender Equality in Party Groups," JCMS: Journal of Common Market Studies, vol. 57, (4), pp. 768-786, 2019.

[47] R. Salahodjaev and D. Jarilkapova, "Female parliamentarism and genuine savings: A cross-country test," Sustainable Development (Bradford, West Yorkshire, England), vol. 27, (4), pp. 637-646, 2019.

[48] A. Altuzarra, C. Gálvez-Gálvez and A. González-Flores, "Is gender inequality a barrier to economic growth? A panel data analysis of developing countries," Sustainability (Basel, Switzerland), vol. 13, (1), pp. 1-21, 2021.

[49] V. Dimitrova-Grajzl and I. Obasanjo, "Do parliamentary gender quotas decrease gender inequality? The case of African countries," Constitutional Political Economy, vol. 30, (2), pp. 149-176, 2019.

[50] ILO, 2020, https://ilostat.ilo.org/data/   (date of Access: 1 May 2022)

[51] M. Fad'os and M. Bohdalova, "Unemployment gender inequality: evidence from the 27 European Union countries," Eurasian Economic Review, vol. 9, (3), pp. 349-371, 2019.

[52] H. H. Dang and C. Viet Nguyen, "Gender inequality during the COVID-19 pandemic: Income, expenditure, savings, and job loss," World Development, vol. 140, pp. 105296-105296, 2021;2020.

[53] C. Shi et al, "A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm," EURASIP Journal on Wireless Communications and Networking, vol. 2021, (1), pp. 1-16, 2021.

[54] C. Subbalakshmi et al, "A Method to Find Optimum Number of Clusters Based on Fuzzy Silhouette on Dynamic Data Set," Proceedings of the International Conference on Information and Communication Technologies, Icict 2014, vol. 46, pp. 346-353, 2015.

[55] L. Lei, S. Desai and R. Vanneman, "The Impact of Transportation Infrastructure on Women's Employment in India," Feminist Economics, vol. 25, (4), pp. 94-125, 2019.

[56] R. P. Pande, S. Namy and A. Malhotra, "The Demographic Transition and Women's Economic Participation in Tamil Nadu, India: A Historical Case Study," Feminist Economics, vol. 26, (1), pp. 179-207, 2020.

[57] M. W. Berhe, "Empirical analysis of urban youth unemployment in Ethiopia," African Development Review, vol. 33, (1), pp. 104-116, 2021.