

Analysis of Well-Being for OECD Countries by Data Mining Methods

By

Seyyit Enes Yurcu

**A Thesis Report Submitted to the
Faculty of Engineering in Partial Fulfilment of the
Requirements for the Degree of**

BACHELOR OF SCIENCE

(Academic Project)

Department: Computer Engineering

**İzmir Institute of Technology
İzmir, Turkey**

June, 2023

We approve the thesis of **Seyyit Enes Yurcu**

Date of Signature

.....
Buket Öksüzoglu

Supervisor

Department of Computer Engineering

Table of content

1. INTRODUCTION.....	3
2. LITRATURE SEARCH RESULTS AND RELATED WORKS	3
3. RESEARCH METHODOLOGY.....	5
4. EXPERIMENTAL SETUP AND EVALUATION.....	7
1. Civic Engagement	9
2. Environmental Quality	12
3. Health.....	20
4. Housing	34
5. Income and Wealth (Finance)	41
6. Knowledge and Skills (Education).....	52
7. Safety (Security)	72
8. Social Connections	77
9. Work and Job Quality	81
10. Work Life Balance.....	109
5. CONCLUSION & IMPACT	112
6. DISCUSSION & FUTURE WORK.....	114
7. REFERENCES.....	115
8. APPENDIX.....	119
A. APPENDIX A: GANTT CHART	119
B. APPENDIX B: ALL MATERIALS.....	119

1. INTRODUCTION

Well-being of countries can be defined as the welfare of the citizens in that country. Financial and social statistics alone do not provide enough information to compare this level. Well-being level also depends on factors such as Health, Education, Housing, Job Opportunity, Social Relations, Safety etc.

The Organization for Economic Cooperation and Development (OECD) provides us with a robust database to understand this level of welfare and the factors on which it depends. OECD has already provided a report about welfare, but this doesn't give clean results and not enough to detect similar countries.

In this academic project, we will seek answers to questions such as "What is the relationship between the health system and the level of well-being of the country?", "How income and wealth distribution affects the well-being of the country?", "How social status affects the well-being of the country?", "Which countries are similar in terms of economy?"etc.

With this project, we will provide a report for the current well-being status of countries and our predictions based on the data.

We will highly use Data Mining methods to analyze the relations of factors and predict future status of countries.

2. LITRATURE SEARCH RESULTS AND RELATED WORKS

These are the projects that similar to our project:

ANALYSIS OF OECD HEALTH DATA WITH DATA MINING METHODS

The research was carried out as a master's thesis by Kübra Coşar(who was a student in Marmara University) in 2020.

The research is mainly focused on Health sector of OECD countries. Based on the data between 2011-2016 years, she developed a data model using Data Mining techniques. The project has no predictive analysis.

We have used similar Data Mining techniques, but have focused on larger topic. So that, we amde analysis of countries much more detailed and in a larger perspective. Also, with our dataset, we were able to make predictions for countries when we had enough data(e.g., which country is more likely to increase their GDP in first 3 upcoming years).

HOW SHOULD WE MEASURE THE “ECONOMIC” ASPECTS OF WELL-BEING?

The research was conducted by CSLS Founder and Executive Director Andrew Sharpe and Economist Lars Osberg on behalf of the Center for the Studies of Living Standards (CSLS) in 2010.

The research contains information about how we should measure wellbeing for Economical part. They have explained well about the economics with charts.

Our project provides more aspects, not only Economics. Also, we have commented more about the data that we analyzed. We used Data Mining methods to analyze, predict which they did not use.

ANALYSIS OF WELL-BEING IN OECD COUNTRIES THROUGH STATIS METHODOLOGY

The research was carried out by F. J. Rivadeneira, A. M. S. Figueiredo, F. O. S. Figueiredo, S. M. Carvajal and R. A. Rivadeneira who were the academicians at universities which was located in Ecuador and Portugal and the research was completed in 2016.

The research is mainly focused on Statistical Analysis of Well-being and factors. So, the research is mostly done by mathematical calculations and compared the countries with each other. They mostly used Euclidean Distance to measure difference between values. Then they have created a trajectory for each country in charts. So, they mostly considered the Statistical and Mathematical part of the data.

This research is already 6 years old. We have provided a new look with fresh data, and a different perspective.

In our research, we analyzed similar OECD data using Data Mining intensively.

In addition, we not only analyzed historical data, but also made predictions for the future when data were available.

APPLICATION OF DATA MINING IN MEDICAL SCIENCE AND HEALTH CARE

The research was carried out as a master's thesis by Didem ATİKTÜRK TAŞDELEN (who was a student in Ankara University) in 2019.

The research is mainly focused on usage of Data Mining techniques to the Health Sector. The dataset is taken from a hospital. The project is about making a data model of patients. With this model, it can make predictions.

Our project is mainly different. We have used Data Mining techniques too, but we focused on different topics.

MOVING FROM A GDP-BASED TO A WELL-BEING BASED METRIC OF ECONOMIC PERFORMANCE AND SOCIAL PROGRESS: RESULTS FROM THE INDEX OF ECONOMIC WELL-BEING FOR OECD COUNTRIES, 1980-2009

The research was conducted by CSLS Founder and Executive Director Andrew Sharpe and Economist Lars Osberg on behalf of the Center for the Studies of Living Standards (CSLS) in 2011.

The research is based on only Economic and Social part of Well-Being. They have provided data, charts, graphs but there is no Data Mining in this project. They have made comments about these data, ideas for future.

Our project has provided more aspects such as Health, Social, Safety instead of starting from 1980s, we have concentrated on more fresh data (after 2005). We used Data Mining methods to analyze, predict about the data collected. Also, the research is completed in 2010 and newest data was in 2009. We have provided a fresh look for countries with every aspect of well-being.

3. RESEARCH METHODOLOGY

Cosine Similarity

Cosine similarity is a similarity metric based on proportionally comparing the increase and decrease similarity of two series of numbers.

It is the cosine of the angle between the vectors.

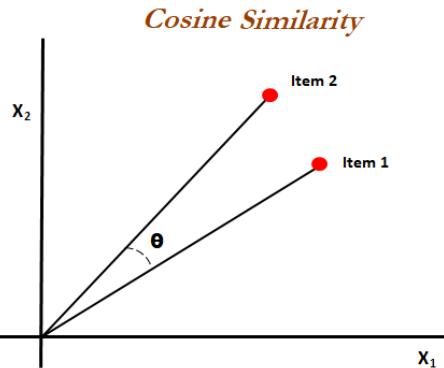


Figure 1: Cosine Similarity of 2 Points

Note. Cosine Similarity. Cosine Similarity From statistics for Machine Learning, n.d. (<https://www.oreilly.com/library/view/statistics-for-machine/9781788295758/eb9cd609-e44a-40a2-9c3af16fc4f5289a.xhtml>).

$$\text{cosine similarity} = \cos(\theta) = (\mathbf{a} \cdot \mathbf{b}) / (\|\mathbf{a}\| \times \|\mathbf{b}\|)$$

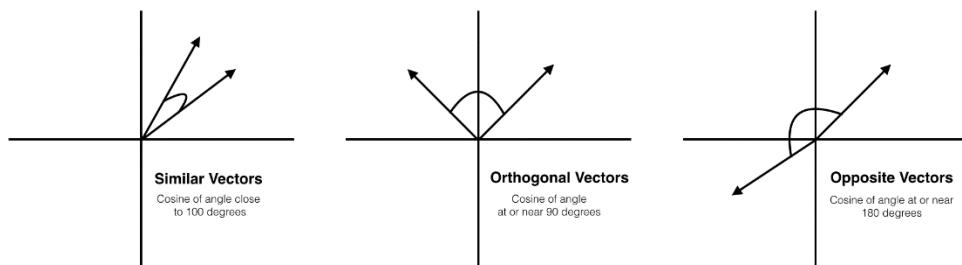


Figure 2: Similarity of Vectors

Note. Vector Similarity. Cosine Similarity From statistics for Machine Learning, n.d. (<https://www.oreilly.com/library/view/mastering-machine-learning/9781785283451/ba8bef27-953e-42a4-8180-cea152af8118.xhtml>).

Cosine Similarity = -1 means exactly opposite vectors

Cosine Similarity = 1 means exactly the same vectors

Cosine Similarity = 0 means indicating orthogonal vectors

Cosine Similarity is an algorithm mostly used for recommendation systems for detecting similar texts by similar "letters" in "words". Cosine similarity measures the cosine of the angle in a multi-dimensional space. We used this algorithm to measure similar increase and decreases in number arrays.

For example, let $a = [10, 20, 30, 40]$, $b = [100, 200, 300, 400]$. In this case, the cosine similarity value is 1.

Similarity of 2 value

For the similarity of two numerical values, we developed the following formula:
similarity= $1 - |(|value1 - value2| / (value1 + value2))|$

Thanks to this formula, it can go up to 1 if two values are equal, and 0 if they are not. The only requirement is that both values must be positive or negative. Otherwise, the similarity value obtained will be negative.

The only requirement is that both values must be positive or negative. Otherwise, the similarity value obtained will be negative which is out of range.

This method is mainly used for datasets with single year, but it is also used for datasets with multiple years based on average similarity of these 2 datasets.

Clustering

Affinity propagation is a clustering method used to group similar data.
In this algorithm, the number of clusters or centroids is not determined as in k-means and k-medoids. Therefore, it is an algorithm that is not constrained by two parameters.

The algorithm groups data with high similarity based on the NxN similarity matrix created with the N-length dataset.

Briefly, this algorithm clusters the data with high similarity according to the predetermined similarity matrix given as input.

Polynomial Regression

Example polynomial: $y=a.x^n+b.x^{n-1}+\dots+e.x+f$

A polynomial is an "n-th" degree formula based on some variables and constants in X axis.
The purpose of polynomial regression is to find a polynomial that fits the data best. The found polynomial is used by plotting to estimate undetermined data in 2-dimensional space.

One of the methods used to measure the reliability of the polynomial obtained by polynomial regression is the " R^2 score". The R^2 score is based on the difference between the y-value obtained by substituting the value for x in the polynomial and the actual y-value. R^2 score is between 0 and 1 and closer R^2 score to 1 is more consistent polynomial.

Datasets

We retrieved all datasets from OECD as “.csv” files. Every dataset is mentioned in detail at the beginning of each section.

The datasets in our project are of 2 types: One is the data for a single year and the other is the data containing changes according to the years.

1-Dataset with a single year

For a single year's data, we determined the similarity ratio of the values using our own similarity formula. Cosine Similarity is not enough to compare a single value.

2-Data with multiple years

In data containing changes by years, we have always taken the common years with data as a basis.

Let's say country A has data for the years 2000,2002,2004,2006 and country B has data for the years 2000,2003,2006. We have taken the common years, that is, the years 2000, 2006, in which both countries contain data, as a basis.

If there are no years in common for two countries, we fill in the gap years with the average values suitable for that country and then get the common years.

If there is still no common year, we made a comparison according to the general average values of the countries.

a-Dependent on Constant Data

If the data we are comparing is not dependent on a constant such as percentage, we tried to capture increasing and decreasing similarity with Cosine similarity.

b-Not Dependent on Constant Data

If the values we compare are dependent on a constant such as percentage (per 100,000 population, per 10,000 motor vehicles etc.), we compare the values as they are for the years in common using our own similarity formula and get an average similarity value.

With these similarity values, a similarity matrix is formed by comparing each country with each other. Based on this similarity matrix we used " Affinity Propagation " to group countries according to their similarities.

4. EXPERIMENTAL SETUP AND EVALUATION

As we are not always able to find data for all OECD countries, we focused on the following 16 countries where we can find data more easily because they are developed. These countries are: Australia, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, Netherlands, Spain, Sweden, Switzerland, Türkiye, United States of America, Great Britain. and country codes for these countries are:

AUS, BEL, CAN, DNK, FIN, FRA, DEU, ITA, JPN, NLD, ESP, SWE, CHE, TUR, USA, GBR.

Implementation

Python was used as the programming language. "sklearn" library in Python is used for "Affinity Propagation". The "seaborn" library was used to visualize the "Data Matrix", and the "pyplot" in the "matplotlib" library in Python was used to graph the data.

Steps:

1. Similarity matrix is created with cosine similarity or similarity formula that explained in the “Research Methodology” section.
2. The created matrix was clustered using "Affinity Propagation" clustering algorithm that explained in the Research Methodology section.
3. Polynomial Regression is also used when its applicable to predict values using the best degree based on R2 score.
4. Clusters and the resulting data are visualized using the appropriate method.

All these 4 steps are applied to 34 datasets for 16 countries.

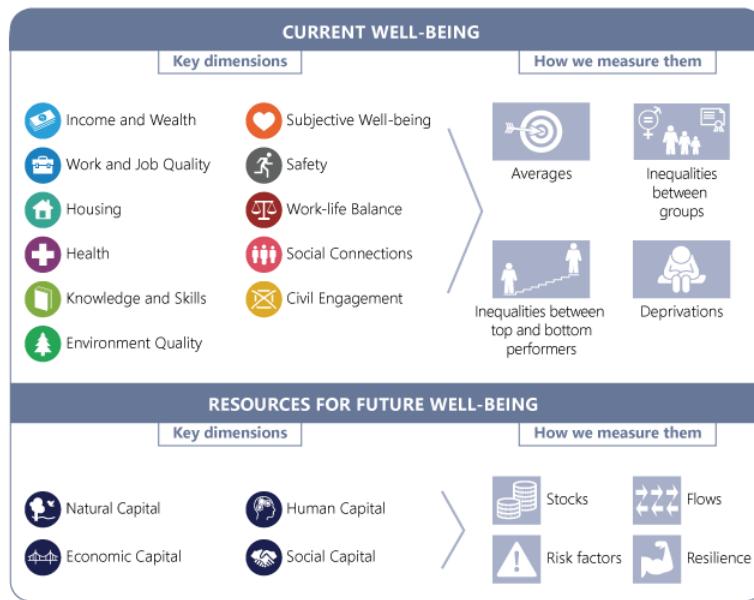


Figure 3: How to measure well-being according to OECD.

Note. OECD well-being framework. From Measuring Well-being and Progress: Well-being Research by OECD, n.d. (<https://www.oecd.org/wise/measuring-well-being-and-progress.htm>).

Figure 3 shows how the OECD measures the current well-being of countries. The main measures of current well-being are based on the economic, health and social conditions of the population. Therefore, we gathered all available datasets on these fields from the OECD. The distribution of the 36 datasets we collected according to their fields is as follows:

1. Civic Engagement
 - a. Civic Engagement
 - b. Voter Turnout
2. Environmental Quality
 - a. Air Pollution
 - b. Environment Politics Index
 - c. Green Growth
3. Health
 - a. Cancer Rate
 - b. Food
 - c. Health Status
 - d. Hospital
 - e. Life Expectancy
4. Housing
 - a. Home Prices
 - b. Home Rent
 - c. Number of Rooms per individual
5. Income and Wealth (Finance)
 - a. GDP per Capita Growth
 - b. Income Inequality
 - c. Exchange Rates
 - d. General Government Deficit
6. Knowledge and Skills (Education)

- a. Foreign Students Rate
 - b. PISA
 - c. Skill Mismatch
 - d. Skill Over Job Quality
 - e. Student per Instructor
7. Safety (Security)
- a. Crime
 - b. Traffic Accident
8. Social Connections
- a. Social Connections
 - b. Social Spending
9. Work and Job Quality
- a. Employment
 - i. Employment Rate by Gender
 - ii. Employment by Education Level
 - b. Unemployment
 - i. Unemployment Rate
 - ii. Unemployment by Education Level
 - iii. Youth Unemployment
 - iv. Long Term Unemployment
 - c. Job
 - i. Job Quality
 - ii. Job Strain
 - d. Gross Domestic Spending on R&D
10. Work Life Balance
- a. Work Life Balance

In the sections below, we explained each field and dataset in detail.

1. Civic Engagement

Civic Engagement is the contribution of citizens to the administration of the country. Civic Engagement is measured by factors such as citizens' participation in elections and educated adults taking an active role in the administration.

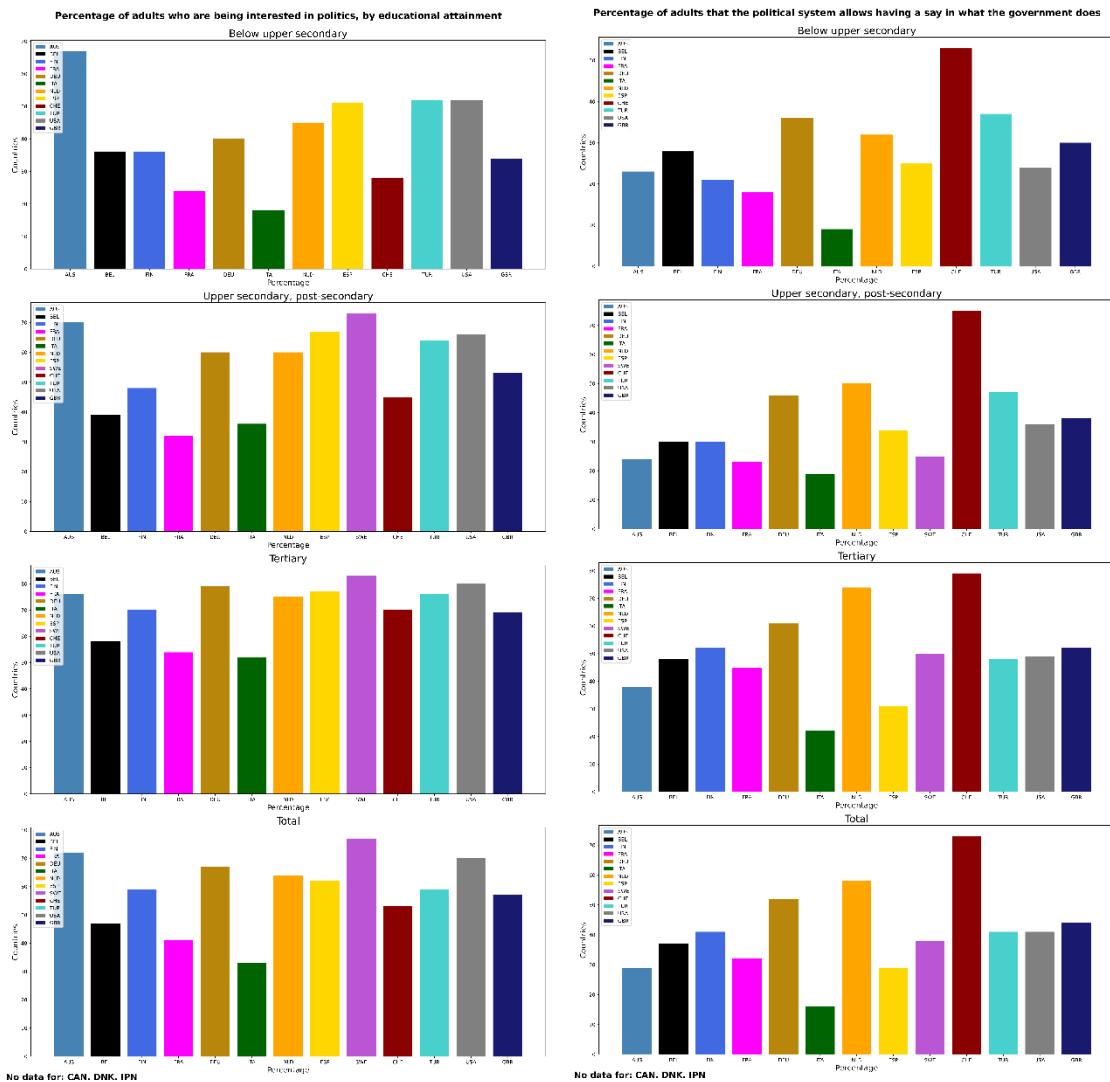
a. Adults in Politics

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Education and Training" section with "Civic engagement and governance" title.

These are all titles in this dataset:

"COUNTRY","Country","ISC11A","ISCED-A
2011","SEX","Gender","AGE","Age","PIAAC_CATEGORY","Category","INDICATOR","Indicator","MEASURE","Measure","YEAR","Reference year","Value","Flag Codes","Flags"

We have 3 educational levels for adults to determine whether "they have a say in what the government does" and "who are being interested in politics".



In terms of interest in politics, Australia is the country with the highest interest among those with a 'below high school' education. France, Italy, and Switzerland show the least interest in this level of education.

As the level of education approaches the highest level, interest in politics increases in France, Italy and especially Switzerland.

When we look at the general situation, interest in politics in Italy is at its lowest level, around 30 percent. When we perform Clustering for the total, we see the following table:

cluster	countries
Cluster 0	['BEL', 'FRA']
Cluster 1	['FIN', 'NLD', 'ESP', 'CHE', 'TUR', 'GBR']
Cluster 2	['ITA']
Cluster 3	['AUS', 'DEU', 'SWE', 'USA']

When we look at the education levels of those who have a say in politics, Switzerland stands out in every education level.

In countries such as the Netherlands, Finland, Germany, and France, as the level of education increases, the rate of having a say in politics also increases.

When we look at the rate of people having a say in politics in total, Italy has the least say and Switzerland has the most say. When we perform Clustering for the total, we see the following table:

cluster	countries
Cluster 0	['ITA']
Cluster 1	['DEU', 'NLD', 'CHE']
Cluster 2	['AUS', 'ESP']
Cluster 3	['BEL', 'FIN', 'FRA', 'SWE', 'TUR', 'USA', 'GBR']

In general, we can say:

- Almost all the educated individuals in Switzerland are involved in politics.
- As the level of education increases, the rate of interest in and having a say in politics increases in Netherlands and Switzerland.
- Although the interest in politics increases as the education level rises in Italy, the rate of having a say in politics does not increase much.

Individuals with lower education have an advantage over more educated individuals. The public more tends to choose those with lower education, because they feel them as "one of the us" and they feel more friendly. (Van Noord et al., 2022) Therefore, it is understandable that educated individuals focus more on their career and academic life. Although these individuals are not interested in politics, states should constantly receive information from them for the future and current status and of the country.

b. Voter Turnout

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the “Regions and Cities” section with “Voters Turnout in regions” title.

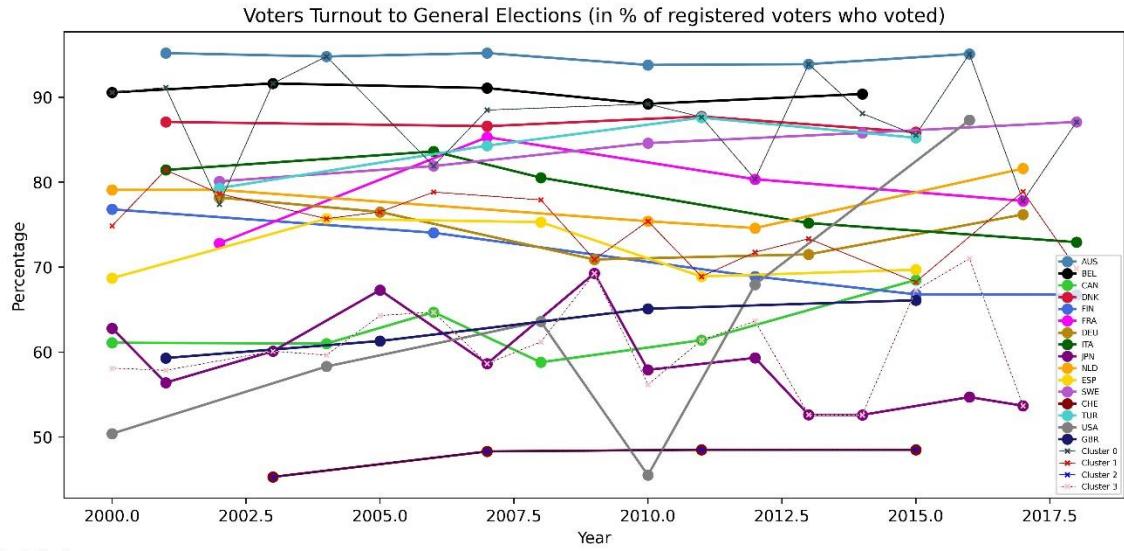
These are all titles in this dataset:

"TL","Territory level and
Typology","REG_ID","Region","VAR","Indicator","POS","Position","TIME","Year",
"Unit Code","Unit","PowerCode Code","PowerCode","Reference Period
Code","Reference Period","Value","Flag Codes","Flags"

Voter turnout is the participation of the people in the government of the country. It is the most democratic right that people choose who wants to come to power. Voter turnout in the elections shows their trust and belief in democracy. This is why voter turnout is very important for democracy.

cluster	countries
Cluster 0	['AUS', 'BEL', 'DNK', 'FRA', 'SWE', 'TUR']
Cluster 1	['FIN', 'DEU', 'ITA', 'NLD', 'ESP']
Cluster 2	['CHE']
Cluster 3	['CAN', 'JPN', 'USA', 'GBR']

Switzerland has its own cluster apart from other countries.



No data for:

- Voter turnout in Switzerland is lower than in other countries and this rate is not likely to increase.
- Australia and Belgium are at the top of the chart with high and constant rate.

As Blais(2016) explains in his research, voter turnout is higher in low-population societies, while in poorer countries it is lower. Looking at the results we obtained in this research, while the voter turnout is lower in countries such as Canada, Japan and the USA, it is higher in Turkey and France, which have similar populations.

According to Blais' research in 2014, this is because the public does not pay due attention to the elections, the civic duty is weak, and the policy system is overly complex. there is an understanding that the "government is a cartel" in Switzerland. Switzerland is governed by a coalition and the results have been pretty much the same in every election for nearly 20 years. The lack of excitement in the election is also a factor that lowers the turnout.

2. Environmental Quality

The environment is very important for human health and well-being. Although it may seem insignificant for some societies, its effects are quite large. For example, air pollution, access to clean water and having a home connected to a sewerage are critical to health.

To see the importance of the environment in society, it is sufficient to look at the importance given to the environment by the administration. Because the rulers of the country are also a part of this great society.

It is never enough for a country to give the necessary importance to the environment in order to increase its environmental quality. Because even though we have political borders on this

planet, we are a whole and we must protect the environment as a whole. Today, we face a threat such as climate change, so it is imperative for countries to work together.

a. Air Pollution

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Regions and Cities" section with "Air pollution" title.

These are all titles in this dataset:

"COU", "Country", "SMALL_SUBNATIONAL_REGION", "Small subnational region", "LARGE_SUBNATIONAL_REGION", "Large subnational region", "VAR", "Variable", "YEA", "Year", "Unit Code", "Unit", "PowerCode Code", "PowerCode", "Reference Period Code", "Reference Period", "Value", "Flag Codes", "Flags"

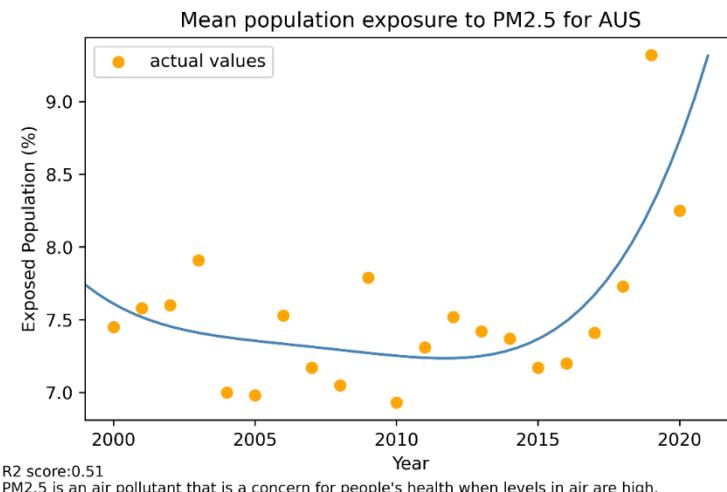
Fine particulate matter PM 2.5 are air pollutants and are very risky for health. Long-term exposure to this matter increases the risk of respiratory and cardiovascular diseases.

When we Cluster the exposure rates to this substance by years according to the similarities of the values, we see the following result:

cluster	countries
Cluster 0	['AUS']
Cluster 1	['CAN', 'USA']
Cluster 2	['BEL', 'FRA', 'DEU', 'ITA', 'NLD', 'ESP', 'CHE', 'GBR']
Cluster 3	['JPN']
Cluster 4	['DNK', 'FIN', 'SWE']
Cluster 5	['TUR']

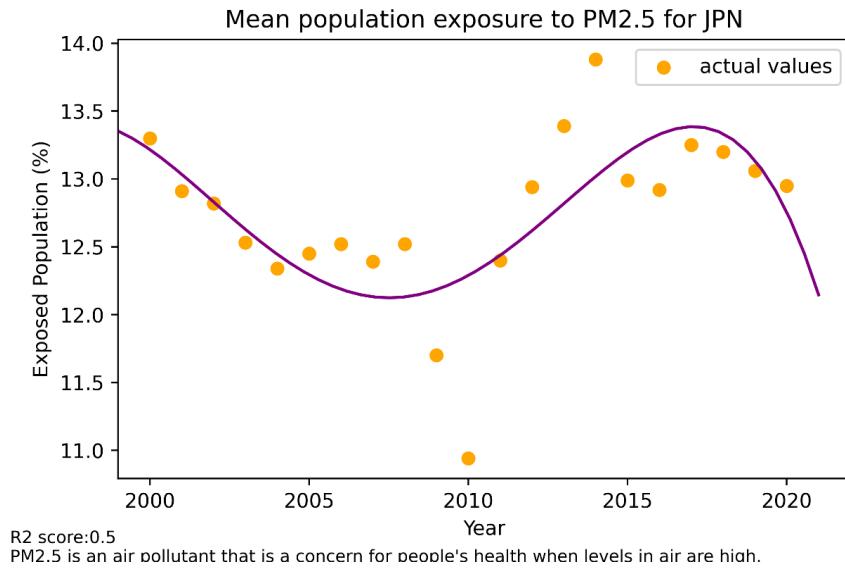
Australia, Japan, and Türkiye are unlike any other cluster. Let's take a closer look at the charts of these countries.

- **Australia**



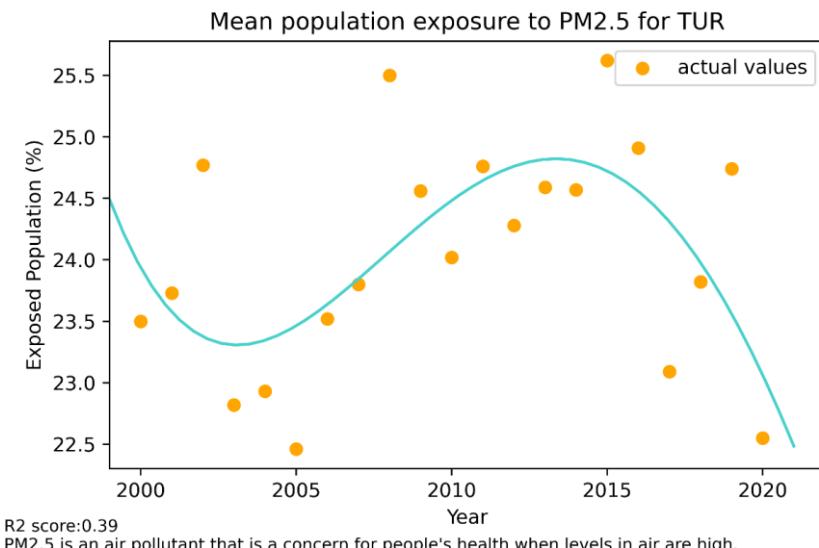
While the exposure rate to PM 2.5 was between 7-8 percent until 2017, then this rate started to increase. The polynomial regression line has R2 score as 0.51, which is not very reliable. Also, by looking at the actual values in the chart, we can say that Australia is not in a stable situation.

- **Japan**



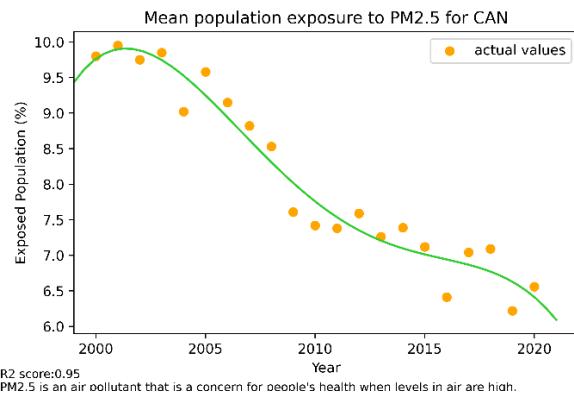
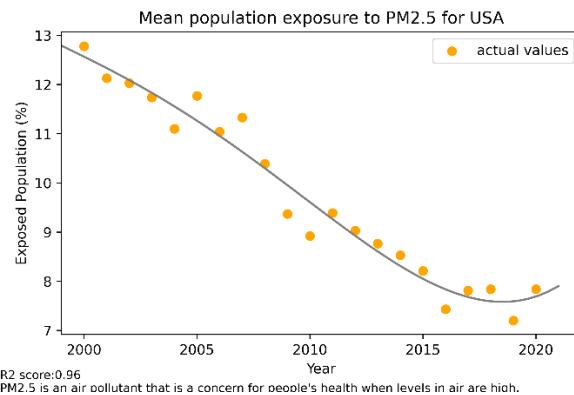
We see a graph similar to a sine wave. Since the R2 score of the polynomial regression line is 0.5 and the actual values on the graph are not very stable, we cannot say that Japan is a very stable country.

- **Türkiye**



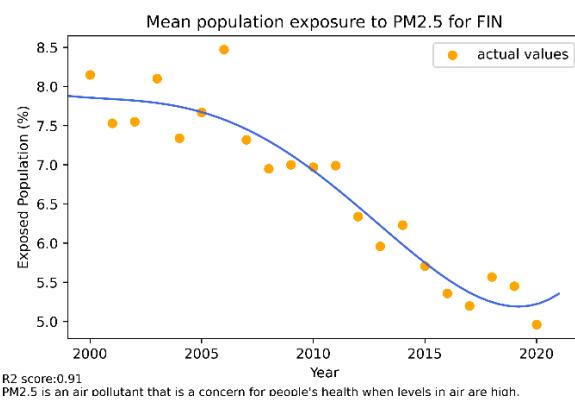
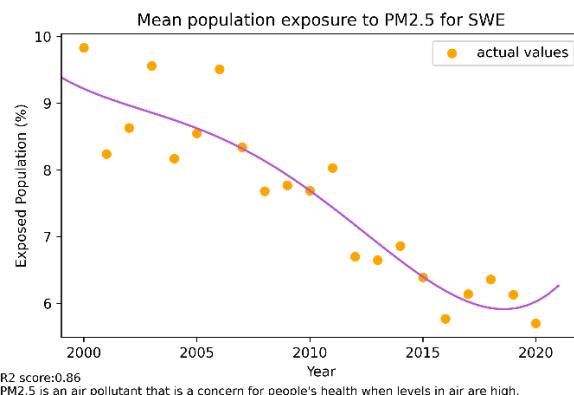
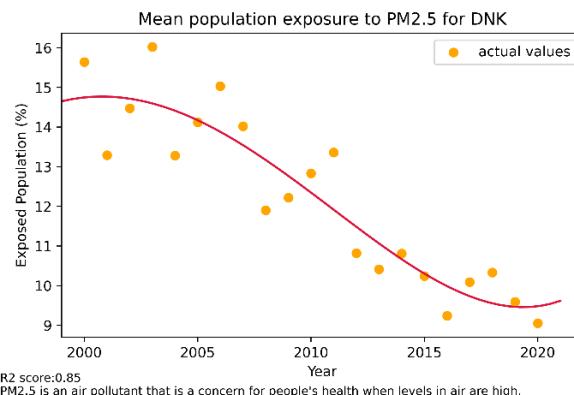
Between 22-25.5, we see very high and very diffuse values. When we try to place these values to a line with polynomial regression, the R2 score of this line is 0.39. R2 score and distribution of actual values shows us Türkiye is very unstable and high percentage of people are exposed with PM 2.5.

- **Canada, USA (Cluster 1)**



These two countries are very similar in values and the ratio is steadily decreasing.

- **Denmark, Finland, Sweden (Cluster 4)**



The PM 2.5 ratio of these 3 Scandinavian countries is steadily decreasing.

Rate of exposure to PM 2.5 is a very effective factor to observe pollution in the air. While some countries have reduced the pm 2.5 exposure with the right policies, some countries have unstable values. This is a big problem that needs to be considered for countries with unstable values such as Türkiye. Because the long-term exposure of PM 2.5 is very dangerous for health and countries must make more strict policies to prevent these health problems (Schlesinger,2007). It is possible to prevent this by applying the right policies.

b. Environment Politics Index

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the “Environment” section with “Environmental Policy Stringency Index” title.

These are all titles in this dataset:

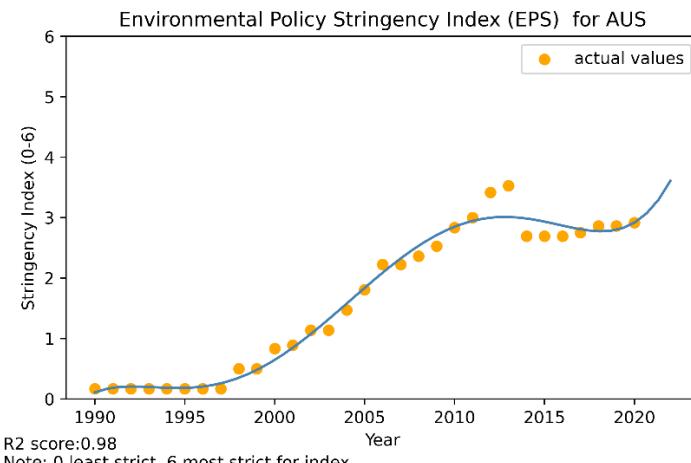
"COU","Country","VAR","Variable","YEA","Year","Unit Code","Unit","PowerCode Code","PowerCode","Reference Period Code","Reference Period","Value","Flag Codes","Flags"

The Environmental Policy Index shows the importance the government attaches to environmental policies. This value is rated from 0 as the loosest, to 6 as the most stringent.

cluster	countries
Cluster 0	['AUS']
Cluster 1	['CAN']
Cluster 2	['FIN', 'JPN', 'CHE']
Cluster 3	['BEL', 'DNK', 'FRA', 'ITA', 'USA', 'GBR']
Cluster 4	['DEU', 'NLD', 'ESP', 'SWE']
Cluster 5	['TUR']

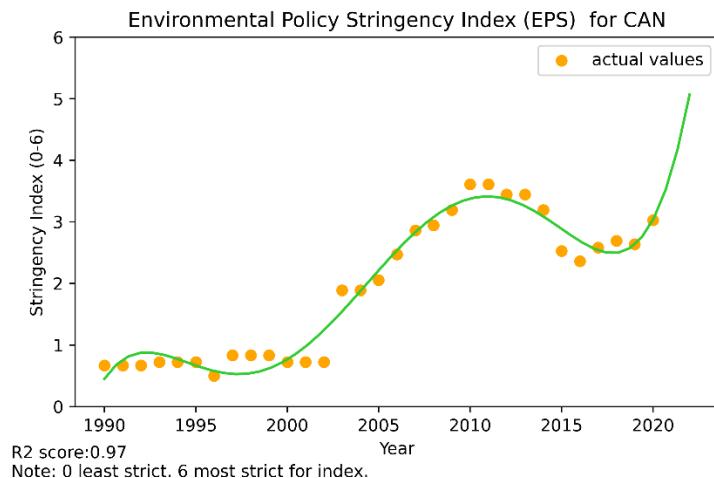
Australia, Canada, and Türkiye are unlike any other cluster. Let's take a closer look at the charts of these countries:

Australia



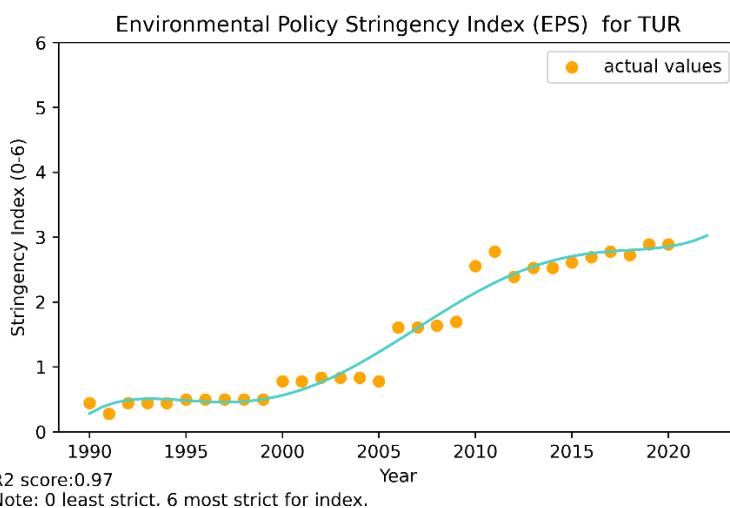
- Australia is steadily tightening its environmental policies. We can expect this index to rise to higher levels.

Canada



- Although Canada has tightened its policies in 2010, it has gone to some relaxation in 2013. However, it would not be wrong to wait for the tightening of policies again.

Türkiye



- Türkiye is tightening its environmental policies very slowly. It would not be wrong to expect this policy to tighten slowly again.

Governments can make people more responsible to the environment with the policies they implement. While countries such as Canada and Australia have been able to achieve this, the Turkish government does not have very strict environmental policies.

The article of Johnson and friends in 2015 also proved that increased policy stringency has a positive impact on environmental development. Therefore, more strict policies can be implemented to increase the environmental consciousness of the public. However, no matter how strict the government's policies may be, ensuring environmental consciousness among the public without providing them with proper education will always remain an insufficient effort.

c. Connected to Sewerage

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Environment" section with "Connection rates to wastewater treatment" title.

These are all titles in this dataset:

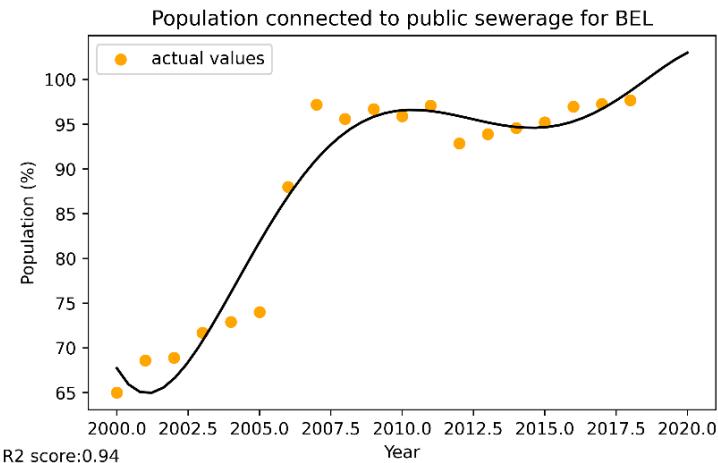
"COU","Country","VAR","Variable","YEA","Year","Unit
Code","Unit","PowerCode Code","PowerCode","Reference Period
Code","Reference Period","Value","Flag Codes","Flags"

Having access to sewage is very important for hygiene and health. For this, it may be sufficient to look at how much of the population in the country is connected to the sewerage in house.

cluster	countries
Cluster 0	['BEL']
Cluster 1	['JPN']
Cluster 2	['TUR']
Cluster 3	['AUS', 'CAN', 'DNK', 'FIN', 'FRA', 'DEU', 'NLD', 'ESP', 'SWE', 'CHE', 'USA', 'GBR']

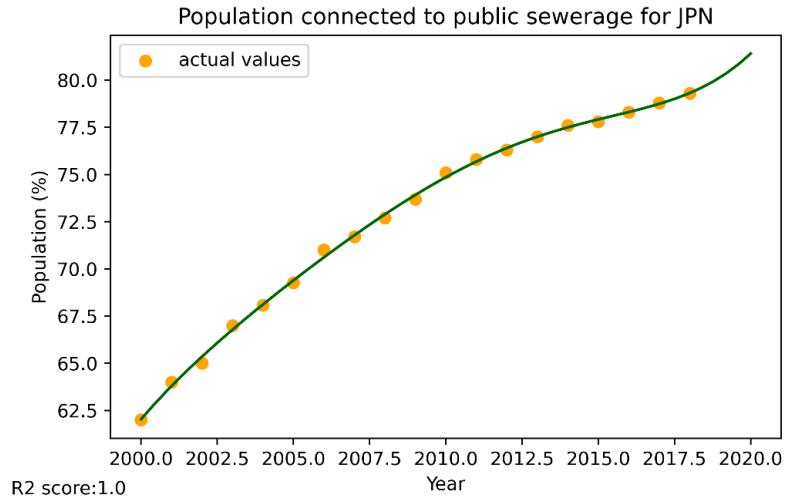
Belgium, Japan, and Türkiye are seemed to be outliers. Let's take a closer look at the charts of these countries:

- **Belgium**



The values suddenly increased after 2005 and reached around 95. Based on successive years of stability and the line drawn by polynomial regression, this ratio can be expected to be close to 100%.

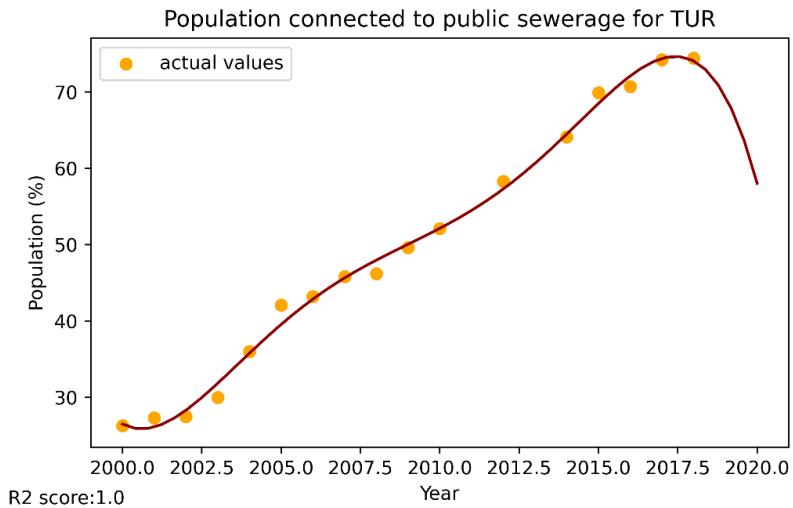
- **Japan**



Japan is steadily and steadily increasing its rate. It would not be a surprise to expect this stability to continue.

However, the reason why a developed country like Japan has such a low access to sewerage needs to be investigated.

- **Türkiye**



Access to sewerage has been steadily increasing from 30 percent in 2000 to the present. This rate is likely to decrease in the future, but this rate should not be that low in a country like Türkiye.

Access to sewerage and clean water is very important for health. Since the immune system is lower in children, it is more important for their health. It is obvious that child mortality rates decreased crucially with the change and development of the city sewer in 1880-1920 (Alsan,2019). Therefore, having a well-developed infrastructure and a proper sewage system is more important than people thought.

3. Health

Good health is the most important thing for a person's life. A healthy life provides more productivity, more income, better relations, less health care costs and most importantly, a longer life.

Life expectancy is one of the most used methods to measure health in a country.

In addition, factors such as the health status of citizens, the number of hospitals in the country, cancer rate and diet also help us to understand the health status of the country.

a. Cancer Rate

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Health" section with "Cancer" title.

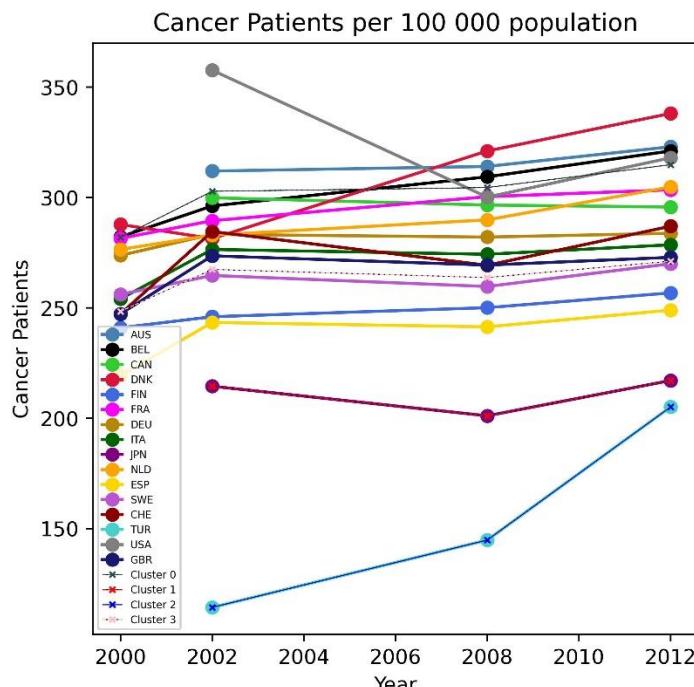
These are all titles in this dataset:

"VAR", "Variable", "UNIT", "Measure", "COU", "Country", "YEA", "Year", "Value", "Flag Codes", "Flags"

Cancer is a very serious disease. Cancer rate in a country can be informative about people with poor health in that country.

cluster	countries
Cluster 0	['AUS', 'BEL', 'CAN', 'DNK', 'FRA', 'NLD', 'USA']
Cluster 1	['JPN']
Cluster 2	['TUR']
Cluster 3	['FIN', 'DEU', 'ITA', 'ESP', 'SWE', 'CHE', 'GBR']

The countries that distinct from out of the two main clusters in this graph are Japan and Türkiye.



The reason why Türkiye and Japan are distinct from these two clusters is that their cancer rates are much lower than in other countries. Türkiye is the country with the lowest cancer rate in the chart.

Although the rates of cancer patients in a country may provide insights into the health-related conditions within that country, it can be expected that people in a community can be more prone to have cancer by factors such as their environment, geography, and lifestyle choices.

Cancer and cancer-related deaths can be prevented through early diagnosis in cancer, cancer control plans organized by the government, and public consciousness campaigns explaining the importance of cancer (Boffetta, 1994). The governments should provide services such as public consciousness campaigns and early diagnosis, especially regarding the types of cancer that is common in the country.

b. Food

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Health" section with "Food supply and consumption" title.

These are all titles in this dataset:

"VAR", "Variable", "UNIT", "Measure", "COU", "Country", "YEA", "Year", "Value", "Flag Codes", "Flags"

The way of diet of the people may depend on the culture, cuisine, and economic reasons of the countries.

For example, if a country is not economically high, the rate of meat consumption is expected to be low. The low rate of meat consumption also affects the amount of protein taken daily. Protein consumption is very important for the human body.

Because protein is used in the construction of cells and development of bone, muscle, and tissue.

1. Sugar Consumption

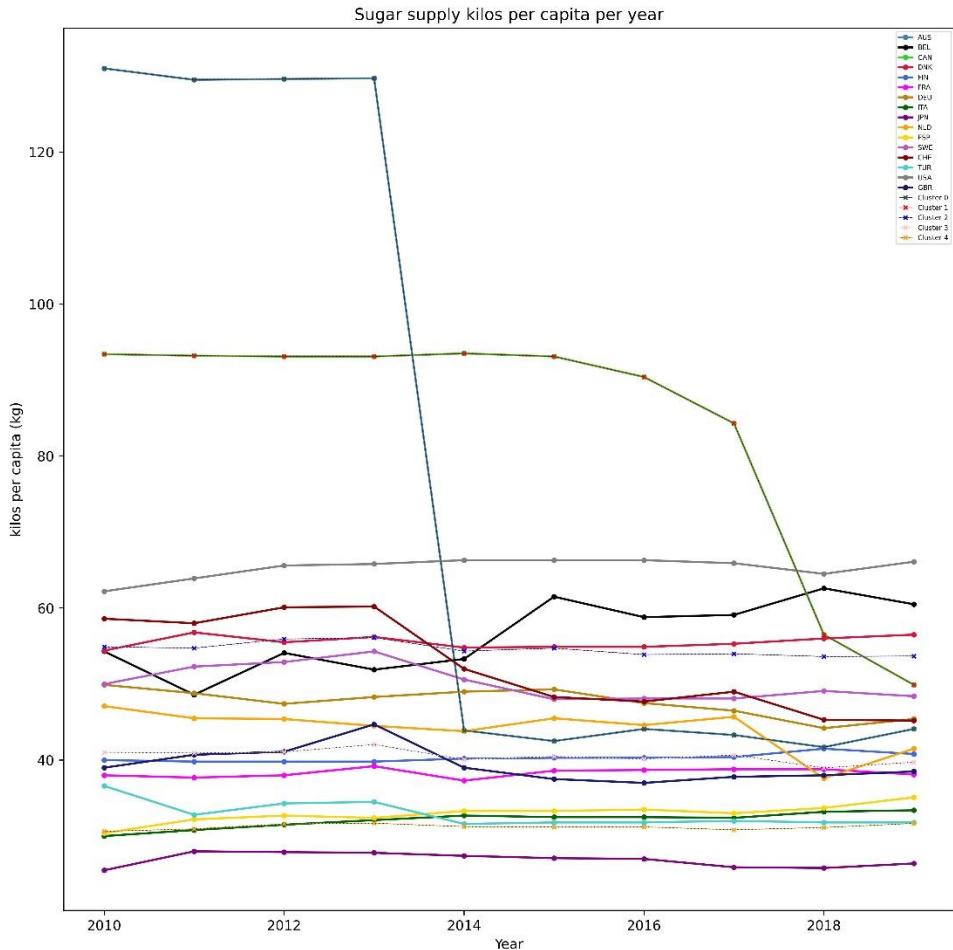
Carbohydrate is the primary energy source. Sugars contain carbohydrates.

Although carbohydrate consumption is important, it is inconvenient to meet the carbohydrate need from sugar. Therefore, the need for carbohydrates must be met from other foods. Sugar can lead to very serious health problems such as obesity, heart diseases and diabetes.

Let's see Clusters of countries with "Sugar Consumption per capita per year":

cluster	countries
Cluster 0	['AUS']
Cluster 1	['CAN']
Cluster 2	['BEL', 'DNK', 'DEU', 'SWE', 'CHE', 'USA']
Cluster 3	['FIN', 'FRA', 'NLD', 'GBR']
Cluster 4	['ITA', 'JPN', 'ESP', 'TUR']

Australia and Canada seem to be alone countries in clusters. Let's take a closer look at the graph:



- While sugar consumption in Australia was extremely high, it started to take place in the middle of other countries with a sharp decline after 2014. It is not surprising that this country is an outlier.
- We can say that there is a noticeable decrease in Canada after 2015. Since there is no other country with a similar picture, it is perfectly logical that it is in a separate cluster on its own.

High consumption of sugar can cause serious health problems. Stanhope clearly explains that sugar directly or indirectly contributes to the emergence of many diseases such as obesity by increasing fat and body mass in her article (2016). Therefore,

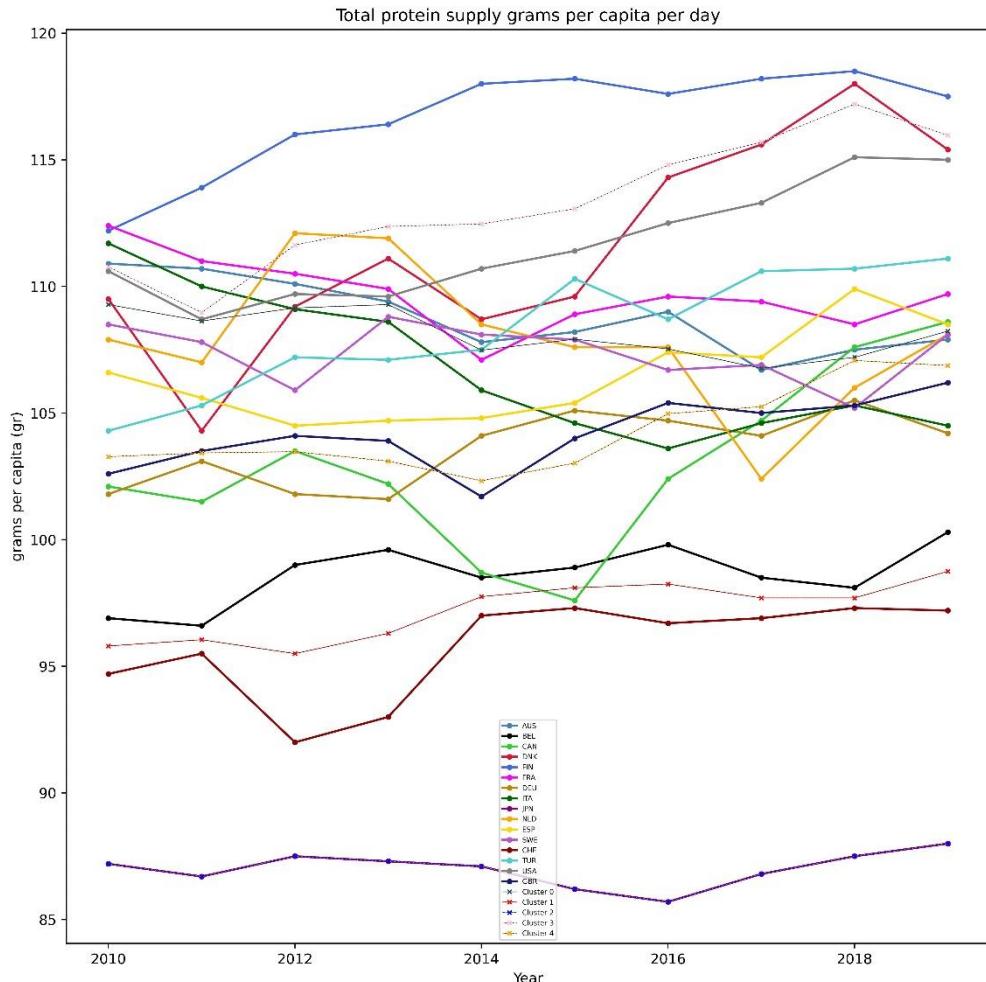
countries with high sugar consumption need to reduce sugar consumption drastically as Australia did. Governments can protect the health of their people with tougher policies, such as raising taxes on products with high sugar.

2. Protein Consumption

Let's see the clusters of "Total protein supply grams per capita per day" in countries:

cluster	countries
Cluster 0	['AUS', 'FRA', 'ITA', 'NLD', 'SWE', 'TUR']
Cluster 1	['BEL', 'CHE']
Cluster 2	['JPN']
Cluster 3	['DNK', 'FIN', 'USA']
Cluster 4	['CAN', 'DEU', 'ESP', 'GBR']

Japan seems to be an alone outlier in the clusters.



- Japan ranks last with low protein consumption. It's perfectly reasonable to be an Outlier.
- Belgium and Switzerland are relatively low compared to other countries and their values are not far from each other. It is quite understandable that these two countries are in the same cluster.

Protein consumption is essential for every age group. Especially the elderly people need to consume more protein. Because the elderly cannot complete the anabolic phase with a low protein intake like the young. In other words, they can only overcome this with higher protein consumption (Baum,2016). Therefore, countries with a high elderly population, especially Japan, need to give much more importance to protein consumption.

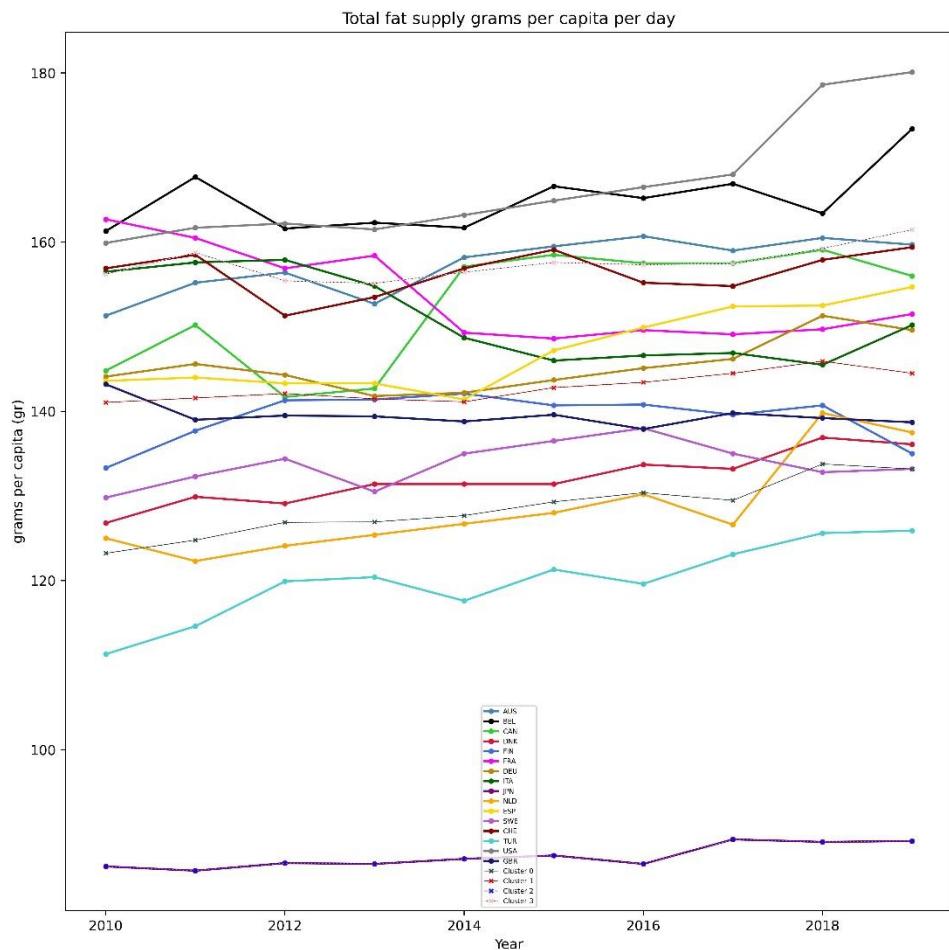
3. Fat Consumption

Fat is essential for the human body. However, the source of this oil must be natural. Eating too much fat is also not healthy. It can lead to very serious problems such as heart attack, fat in the organs.

Also, a country's consumption of too much fat and sugar can be associated with too much fast-food consumption.

cluster	countries
Cluster 0	['DNK', 'NLD', 'SWE', 'TUR']
Cluster 1	['FIN', 'DEU', 'ESP', 'GBR']
Cluster 2	['JPN']
Cluster 3	['AUS', 'BEL', 'CAN', 'FRA', 'ITA', 'CHE', 'USA']

Looking at the cluster, Japan seems to be the outlier. Let's take a closer look at the graph:



- Japan is in the last place with stable values in fat consumption, as in many other foods. It is very reasonable that Japan is an outlier.

Fat consumption is essential for a healthy body and brain. However, people with high fat mass have a higher risk of cardiovascular disease. Also, consuming too much high saturated fat and especially trans-fat results in many cardiovascular diseases (Remig, 2010). Therefore, the amount of fat that should be consumed during the day should be taken from healthy fat sources.

4. Calories per day

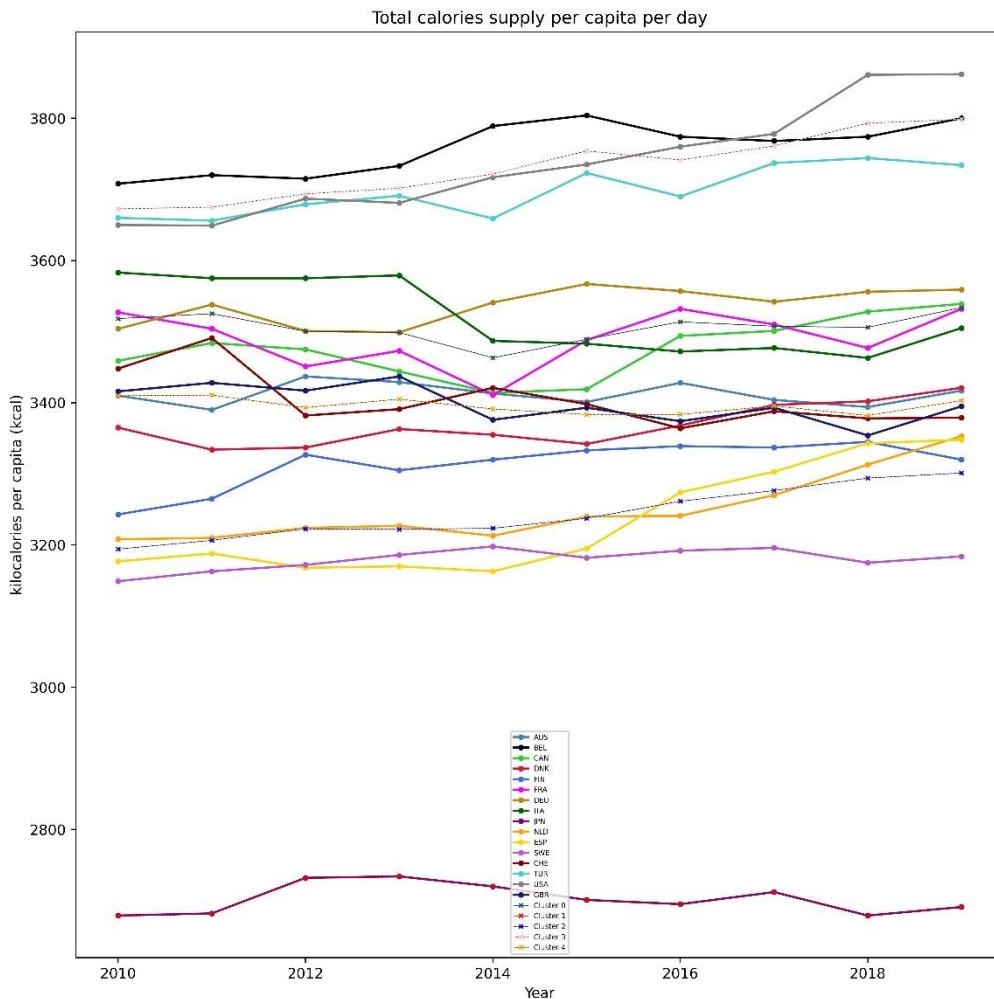
Too much fat consumption increases the amount of calories consumed daily.

Because 1 gram of fat equals to 9 calories, while 1 gram of carbohydrates or 1 gram of protein equals to 4 calories.

Lets take a closer look to the clusters:

cluster	countries
Cluster 0	['CAN', 'FRA', 'DEU', 'ITA']
Cluster 1	['JPN']
Cluster 2	['FIN', 'NLD', 'ESP', 'SWE']
Cluster 3	['BEL', 'TUR', 'USA']
Cluster 4	['AUS', 'DNK', 'CHE', 'GBR']

- Since Japan is very low in protein, fat, and sugar consumption, it makes sense that the total calorie supply is very low. Therefore, it is inevitable that Japan would be an outlier.



- We see that countries that consume a lot of fat, such as the USA and Belgium, rank at the top in total calorie supply. These countries consume a lot of fat as a source of calories, and this is not healthy.
- Although fat consumption is not very high in Turkey, it ranks high in total calorie supply. This shows that people in Turkey do not consume much fat as an energy source and this situation is quite healthy.

As we expected, societies that consume a lot of fat also have a very high daily calorie intake. It is important not to exceed the required daily calorie amount constantly, as carbohydrates and protein taken in excess are converted into fat. Otherwise, many diseases such as obesity and cardiovascular diseases are inevitable due to the high fat rate. Also, eating much less than necessary, like in Japan, leads to a weak immune system in best case.

5. Vegetable and Fruit Consumption

a. Vegetable Consumption

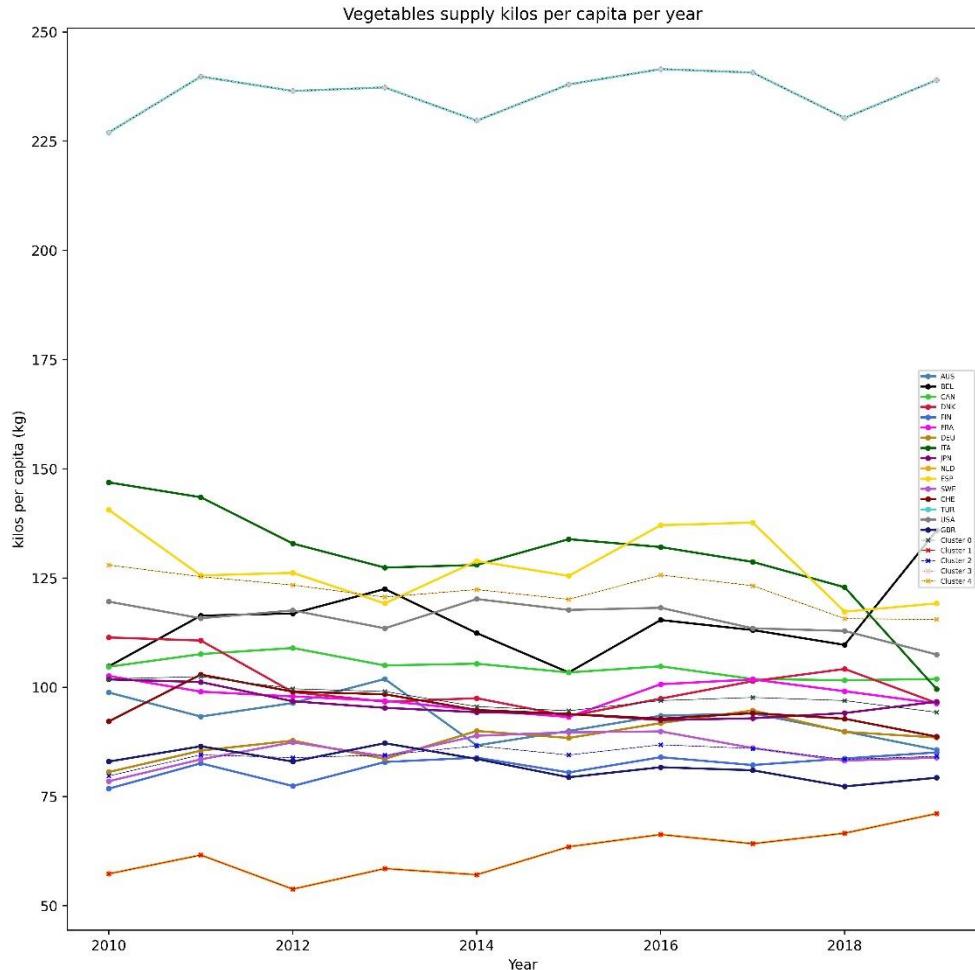
Vegetable consumption is essential for a healthy life. Some cuisines, such as Mediterranean cuisine, have too many dishes prepared with vegetables.

Vegetable consumption is expected to be high in these countries. However, the high consumption of vegetables may also be due to economic reasons.

Let's see the Clusters:

cluster	countries
Cluster 0	['AUS', 'CAN', 'DNK', 'FRA', 'JPN', 'CHE']
Cluster 1	['NLD']
Cluster 2	['FIN', 'DEU', 'SWE', 'GBR']
Cluster 3	['TUR']
Cluster 4	['BEL', 'ITA', 'ESP', 'USA']

Türkiye and Netherlands seem like the outliers in the Clusters. Let's see the graph:



- We can attribute the high consumption of vegetables in Turkey to the Mediterranean cuisine, and it is quite healthy. Turkey consumes twice as much vegetables as the second closest country in terms of vegetable consumption. So, they deserve to be the outlier.
- The Netherlands has a much less consumption than other countries. It is understandable that they are just another outlier in the graph.

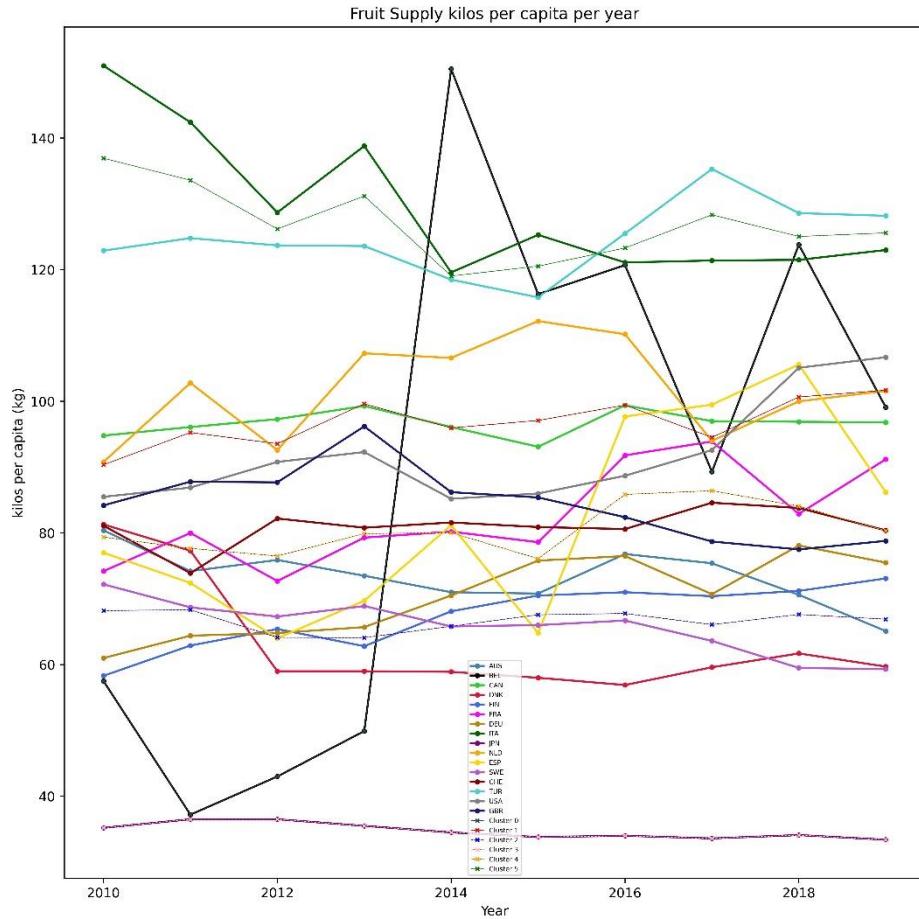
b. Fruit Consumption

Fruits are very healthy foods because of the vitamins and minerals they contain. It would not be wrong to expect a society that consumes fruit to be healthy.

Lets see clusters of Fruit Consumption of countries:

cluster	countries
Cluster 0	['BEL']
Cluster 1	['CAN', 'NLD', 'USA']
Cluster 2	['DNK', 'FIN', 'DEU', 'SWE']
Cluster 3	['JPN']
Cluster 4	['AUS', 'FRA', 'ESP', 'CHE', 'GBR']
Cluster 5	['ITA', 'TUR']

Belgium and Japan seem to be outliers in these clusters. Italy and Türkiye are in the same cluster.



- Annual fruit consumption in Japan is quite low and stable. Since there is no other country that is so low and stable, we can expect that this country is an outlier.
- Belgium's chart is quite different. While fruit consumption was very low at the beginning of the 2010s, this number increased considerably afterwards. Then, we see that there was a decrease again. So, we can say that Belgium was unstable. Since there is no other country with similar values to Belgium's values, it is not surprising that this country is an outlier.
- Italy and Türkiye are two Mediterranean countries. It is not surprising that their diets are similar. These countries are at the top and their values are not very different from each other. It is quite reasonable that the algorithm has put these two countries in the same cluster.

Vitamins and minerals are essential for a healthy body. Since vegetables and fruits contain lots of vitamins and minerals, they have many benefits such as strengthening the immune system and cleansing the blood. It is known that regular consumption of vegetables and fruits reduces the rate of having cancer (Terry, 2001). Since all minerals and vitamins cannot be stored in the body, vegetables and fruits should be consumed every day.

c. Health Status

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Health" section with "Perceived health status" title.

These are all titles in this dataset:

"VAR", "Variable", "UNIT", "Measure", "COU", "Country", "YEA", "Year", "Value", "Flag Codes", "Flags"

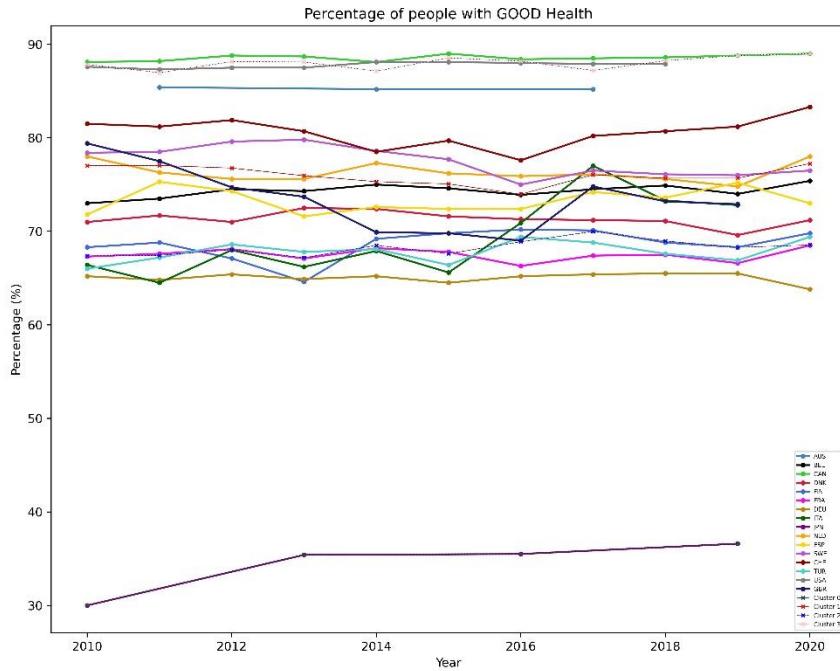
This dataset contains information about the health status of people in the country. There are 3 different tables: Percentage of People with good health, fair health, and bad health.

When we looked at their diet, we saw that Japan consumes much less food than other countries. Therefore, we can expect Japan to have a lower level of health than other countries and to rank alone in many clusters.

1. Good Health

cluster	countries
Cluster 0	['JPN']
Cluster 1	['BEL', 'NLD', 'ESP', 'SWE', 'CHE', 'GBR']
Cluster 2	['DNK', 'FIN', 'FRA', 'DEU', 'ITA', 'TUR']
Cluster 3	['AUS', 'CAN', 'USA']

For the reasons we've mentioned, we only see Japan as an outlier.

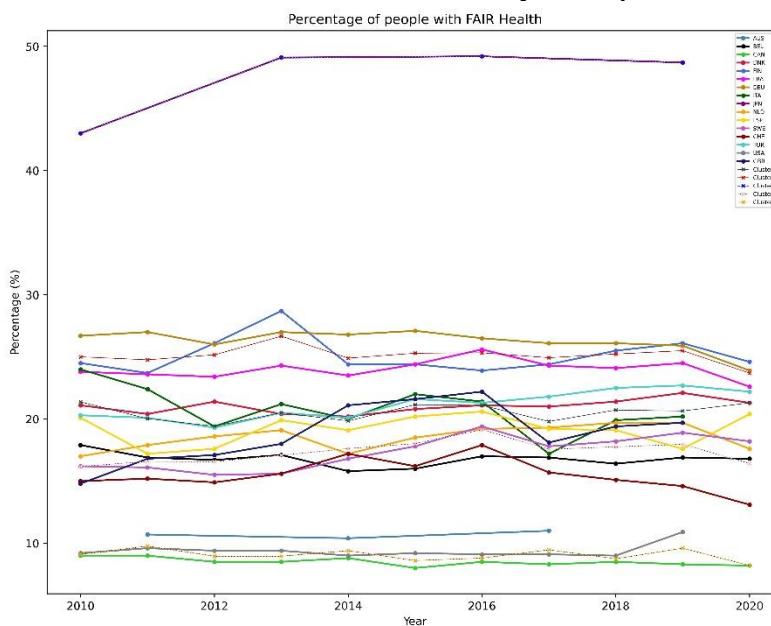


- Australia, Canada, and USA (Cluster 3) have the best percentage of people with good health.

2. Fair Health

cluster	countries
Cluster 0	['DNK', 'ITA', 'ESP', 'TUR']
Cluster 1	['FIN', 'FRA', 'DEU']
Cluster 2	['JPN']
Cluster 3	['BEL', 'NLD', 'SWE', 'CHE', 'GBR']
Cluster 4	['AUS', 'CAN', 'USA']

For the reasons we've mentioned, we only see Japan as an outlier.

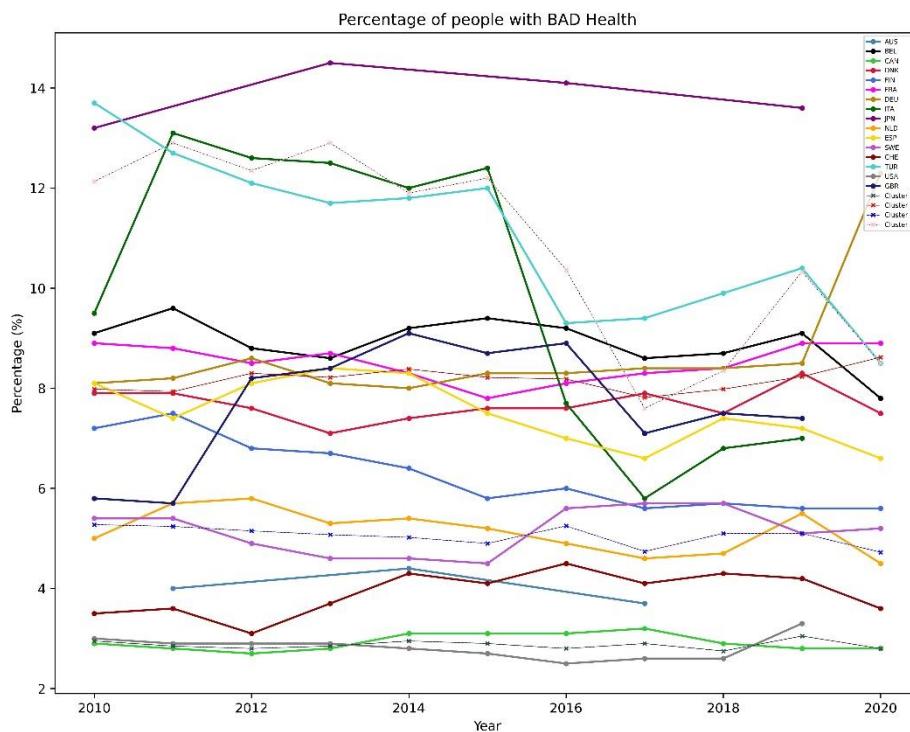


- We see that about half of the people in Japan are in fair health.
- Australia, Canada, and the USA (Cluster 4) are the countries with the lowest percentages of people in fair health.

3. Bad Health

cluster	countries
Cluster 0	['CAN', 'USA']
Cluster 1	['BEL', 'DNK', 'FRA', 'DEU', 'ESP', 'GBR']
Cluster 2	['AUS', 'FIN', 'NLD', 'SWE', 'CHE']
Cluster 3	['ITA', 'JPN', 'TUR']

We see Japan, Italy, and Türkiye in same Cluster. Also, we see Canada and USA in another Cluster.



- Canada, and USA (Cluster 0) are the countries with the lowest percentages of people in bad health.
- In general, we can say that most of the people in Japan have a bad and fair health and most of the people in USA and Canada has a good health. Also, Australia has a pretty good health.

Although this data is survey-based, it is sufficient for us to comment on the health status of people in the country. It is clear that people in Japan, with insufficient nutrition, are not in very good health. The USA people are satisfied with their health despite their high-fat diet.

d. Hospital

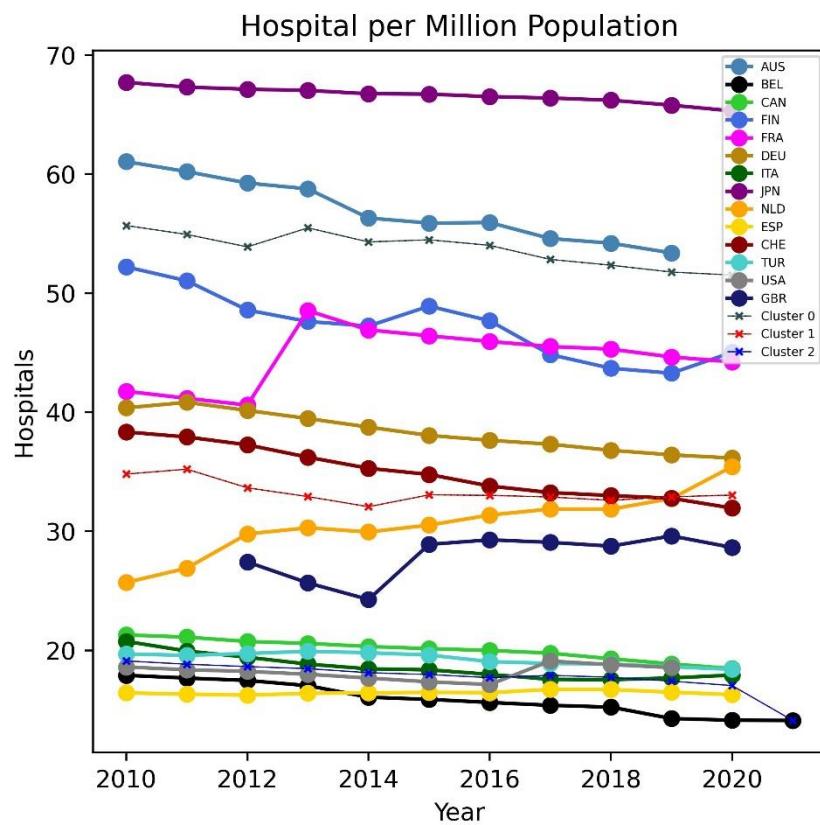
This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Health" section with "Hospitals" title.

These are all titles in this dataset:

"VAR", "Variable", "UNIT", "Measure", "COU", "Country", "YEA", "Year", "Value", "Flag Codes", "Flags"

Hospitals are indispensable institutions of the health system where people find healing and regain their health. Therefore, the number of hospitals per person shows how easily people can reach doctors.

cluster	countries
Cluster 0	['AUS', 'FIN', 'FRA', 'JPN']
Cluster 1	['DEU', 'NLD', 'CHE', 'GBR']
Cluster 2	['BEL', 'CAN', 'ITA', 'ESP', 'TUR', 'USA']



- Australia, Finland, France, and Japan rank highest in the number of hospitals per million population, and this is the worst cluster.
- Germany, Netherlands, Switzerland, and Great Britain are in the middle cluster.
- Belgium, Canada, Italy, Spain, Turkey, and the USA are among the lowest in the number of hospitals per million population and this is the best cluster.

Looking at the table and data above, we can say that the number of hospitals has a very clear effect on the proportion of healthy people.

For example, there is 1 hospital per 70 million people in Japan, and the number of healthy people is quite low.

Likewise, there is 1 hospital per 20 million people in Canada and the USA, and the number of healthy people is quite high.

Policy makers are more concerned with the money in the sector; because many private hospitals are set up for commercial purposes (Varabyova,2013). Therefore, factors such as functionality, number of beds, number of personnel, number of medical equipment are more important than the number of hospitals. Also, rather than the number of hospitals, the quality of hospitals, diet and environmental factors also affect the number of healthy people. For example, although Belgium and Spain are at the top in the number of hospitals per capita, they are not at the top in the percentage of healthy people.

e. Life Expectancy

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Health" section with "Life expectancy" title.

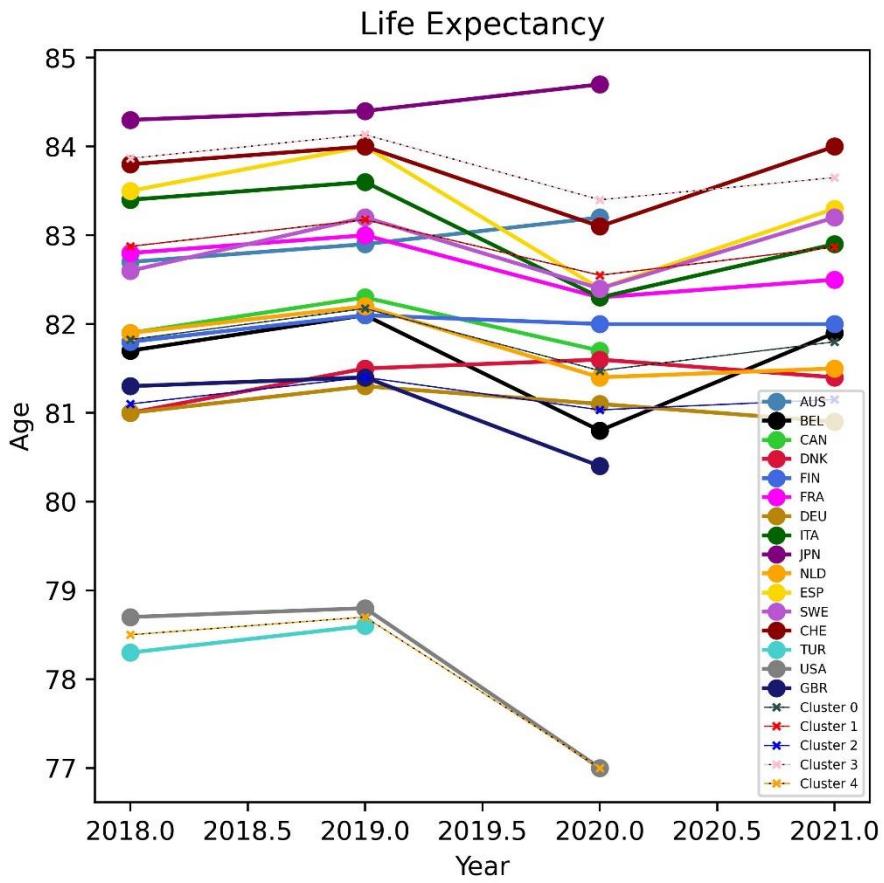
These are all titles in this dataset:

"LOCATION", "INDICATOR", "SUBJECT", "MEASURE", "FREQUENCY", "TIME", "Value", "Flag Codes"

Life expectancy is a measure of how many years a person can expect to live. It gives actual information to understand the health status of people in that country.

cluster	countries
Cluster 0	['BEL', 'CAN', 'FIN', 'NLD']
Cluster 1	['AUS', 'FRA', 'ITA', 'SWE']
Cluster 2	['DNK', 'DEU', 'GBR']
Cluster 3	['JPN', 'ESP', 'CHE']
Cluster 4	['TUR', 'USA']

We see 4 clusters with at least 2 countries for each. So, there is no obvious country with different data. Let's have a closer look to the graph:



- Although the percentage of people with good health is high in the USA and Turkey, life expectancy is quite low.
- Although the percentage of healthy people is low in Japan, life expectancy is quite high.

Although the people of the USA are satisfied with their health status, their life expectancy is quite short compared to other countries. The number of hospitals in Turkey seems to be sufficient, the public is satisfied with their health and eating properly, but the expected life expectancy is quite low. Although, the public is dissatisfied with their health and not enough feeding in Japan, people live much longer than in other countries.

So, we can say that even if people are healthy in a country, there are also many environmental factors such as stress, geography, quality of life and pace etc. So, life is more complex than it seems, and we cannot measure every factor.

Life expectancy has increased by 1.5 times in the last 65 years, with medicine finding cures for diseases and developing technology (Roser,2013). Although the human lifespan has increased, there are thousands of diseases that have no cure, and new diseases will always arise. That's why we should be healthy as human beings and take good care of our body.

4. Housing

Shelter is one of the basic needs of people.

Housing should be a place where people feel safe, where they can sleep, rest, and also have privacy and personal space.

Factors such as house rent, house prices, and the number of rooms per capita help us understand how a country provides housing to its people.

a. Home Prices

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Prices and Purchasing Power Parities" section with "National and Regional House Price Indices" title.

These are all titles in this dataset:

"TL","Territory Level and Typology","REG_ID","Region","VAR","Variable","VINTAGE","Vintage","DWELLINGS","Dwellings","MEASURE","Measure","FREQ","Frequency","TIME","Time","Unit Code","Unit","PowerCode Code","PowerCode","Reference Period Code","Reference Period","Value","Flag Codes","Flags"

Owning a home is one of the first goals of every adult. Because paying rent every month creates a constant expense. For this reason, people want to buy a house as soon as possible. To be able to own a house more easily, house prices and people's salaries should not be too far away.

1 period or 1 quarter means a quarter of a year which is equal to 3 months. We will concentrate on change of prices between quarters and years. We will also consider the change in the price of a house with a price of 100 units in 2015 according to years.

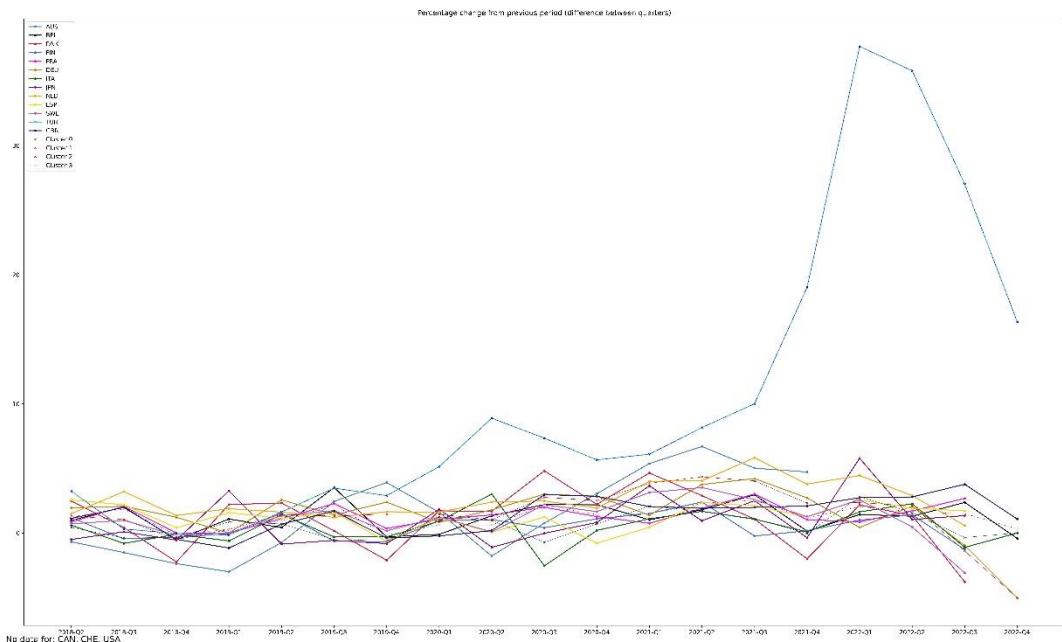
Let's take a closer look at house prices in the countries:

1. Percentage change from previous period

This chart is based on a comparison of the change in the price of the house compared to the previous quarter.

cluster	countries
Cluster 0	['FIN', 'ITA', 'JPN']
Cluster 1	['AUS', 'DNK', 'DEU', 'NLD']
Cluster 2	['TUR']
Cluster 3	['BEL', 'FRA', 'ESP', 'SWE', 'GBR']

We clearly see Türkiye is an alone Cluster. Let's see the graph:



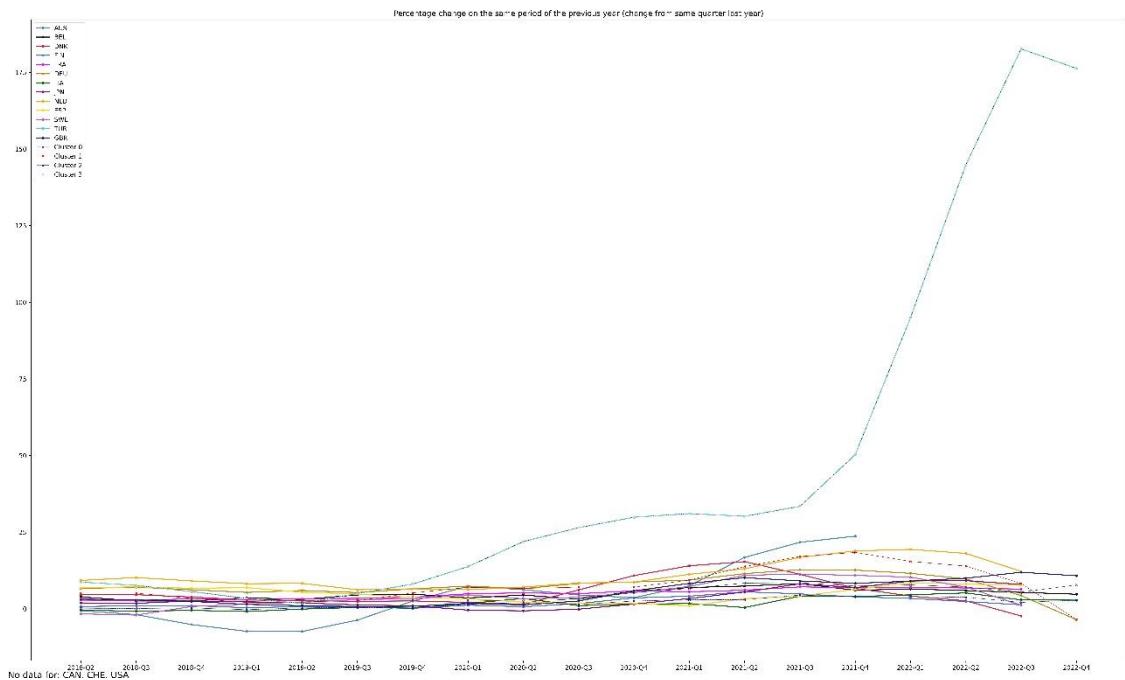
- House prices in Turkey started to increase significantly after 2019. Even if this chart shows the change between quarters, the increase in the last quarters of 2021 is quite high. Although we see a decrease in the rate of increase after 2022, this does not mean that prices have decreased. It just shows that the rate of increase has slowed down a bit. It would be correct to say that Turkey is an outlier because the prices have increased extremely.

2. Percentage change from previous year

This chart is based on a comparison of the change in the price of the house compared to the same quarter in the previous year.

cluster	countries
Cluster 0	['BEL', 'DNK', 'FRA', 'JPN', 'ESP', 'SWE', 'GBR']
Cluster 1	['AUS', 'DEU', 'NLD']
Cluster 2	['FIN', 'ITA']
Cluster 3	['TUR']

We see Türkiye as an outlier which is not surprising.



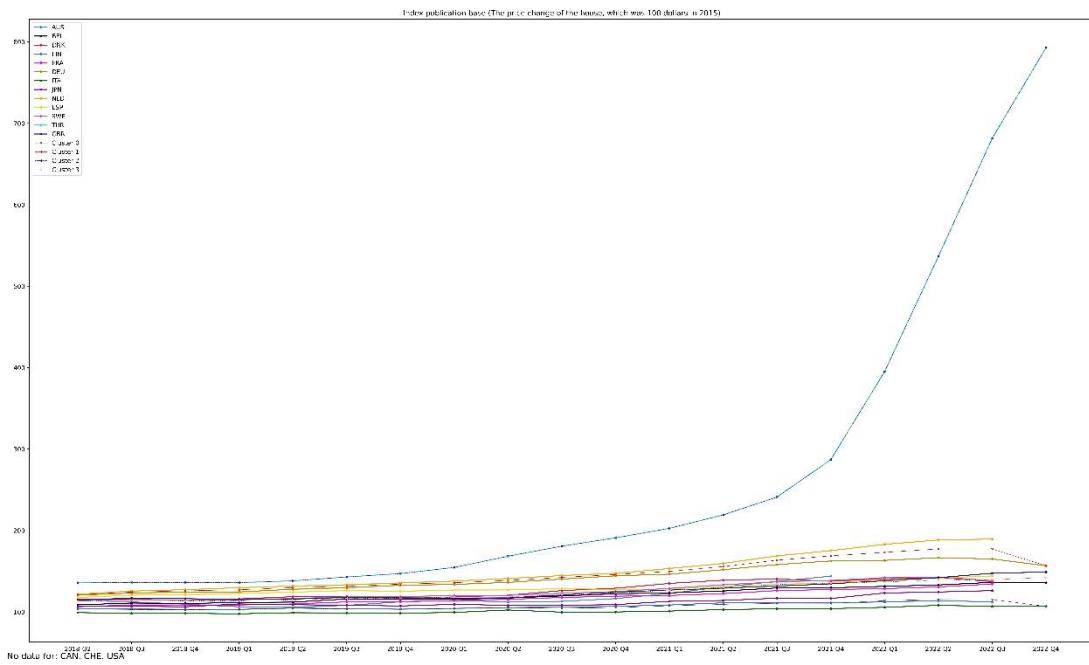
- When we look at Turkey's data, we see that there has been an increase from 100 percent to 175 percent compared to the previous year, especially in the quarters after 2021. In other words, house prices have increased by 2-3 times compared to the previous year. This is very extreme and can make it very difficult for people to buy a house.
- When we look at the graphs of Italy and Finland, we see that these countries are generally at the bottom. Although we see small increases in places, we also see decreases. We can say that the price has remained almost the same in years.

3. House prices based on 2015

We have seen the growth charts by years and quarters. So, what would a 100-unit house cost today in 2015?

cluster	countries
Cluster 0	['FIN', 'ITA', 'JPN']
Cluster 1	['DEU', 'NLD']
Cluster 2	['TUR']
Cluster 3	['AUS', 'BEL', 'DNK', 'FRA', 'ESP', 'SWE', 'GBR']

We see Türkiye as an outlier, this is what we expected. We see Germany and Netherland in same cluster.



- We see that house prices in Turkey are constantly increasing. It is not a common situation that the price of a house which was 100 units in 2015 has reached 8 times in 7 years. As we've just seen from the quarterly and yearly change of home prices charts, we see sharper rises after the end of 2021 in this chart too. This could mean that people spend a lifetime buying a home.
- When we look at the graph of Australia and the Netherlands (Cluster 1), we see a slight increase at the same time and at similar levels. Therefore, we can say that these two countries are quite similar in terms of house prices.
- Finland, Japan, and especially Italy (Cluster 0) are countries where house prices are almost unchanged. It is possible to say that the price of the house, whose price was 100 units in 2015, remained almost the same in these 3 countries. Türkiye should examine the policies in these countries and implement the policies that can be implemented in the country.

According to the supply-demand balance, which is the most basic understanding in economy, the price of the product with less quantity will be high. Therefore, high house prices indicate that there is not enough housing supplied in that country. According to a study conducted in the USA, the reason for the increase in house prices is the low housing supply and the high population in small settlements. Also, New regulations in housing supply are the biggest cause of low housing supply (Glaeser, 2005). So, it is obvious that there is a problem in housing supply in Turkey, where house prices are increasing extremely. In his article he wrote in 2011, Coşkun mentioned that there is a supply and demand gap in the housing sector in Türkiye (Coşkun, 2011). His predictions about the future were correct and the number of houses in Turkey was insufficient. Therefore, house prices have increased and continue to rise.

b. Home Rent

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Regions and Cities" section with "Housing indicators" title.

These are all titles in this dataset:

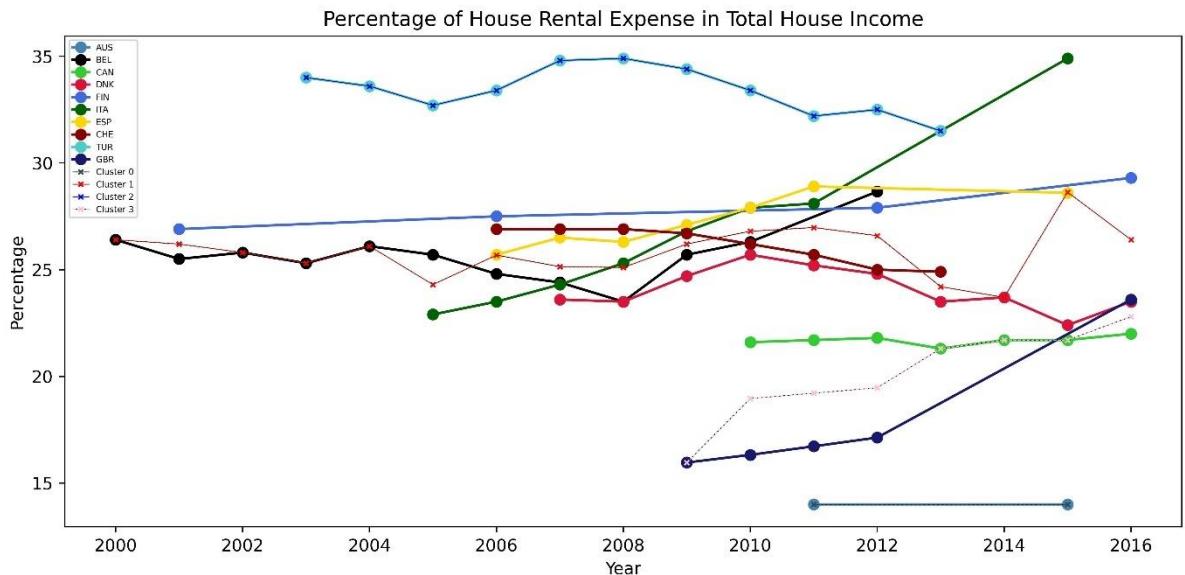
"TL","Territory level and Typology","REG_ID","Region","VAR","Indicator","POS","Position","TIME","Year","Unit Code","Unit","PowerCode Code","PowerCode","Reference Period Code","Reference Period","Value","Flag Codes","Flags"

This dataset is prepared to show what percentage of total household income goes to rent.

As we have seen in house prices, we can expect house rents to be high in Türkiye.

cluster	countries
Cluster 0	['AUS']
Cluster 1	['BEL', 'DNK', 'FIN', 'ITA', 'ESP', 'CHE']
Cluster 2	['TUR']
Cluster 3	['CAN', 'GBR']

We see Türkiye as an outlier which is not surprising. We also see Australia as another alone cluster in this dataset. Let's have a closer look:



No data for: FRA. DEU. IPN. NLD. SWE. USA

- We see that 35 percent of household income in Turkey goes to rent. This is quite high, and Türkiye is at the top in this field. This is not surprising, because home prices in Türkiye is very high and people rent their houses with high prices.
- At the bottom of the chart, we see Australia with 14 percent. So, in Australia people don't spend a lot of money on rent and can spend a lot more on other needs.

Rent is the money paid to the owner because of using a product that you do not own. House rent, which is paid to the owner of house regularly every month, is a permanent expense. For this reason, it is important for people to own a house. In countries where house prices are very high, it becomes more difficult for people to own a house and the number of people living in rent increases. Thus, there is a constant flow of money from the low-income to the high-income and income inequality in the country increases. Therefore, lower house prices are very important to avoid income inequality.

c. Number of Rooms per individual

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the “Regions and Cities” section with “Housing indicators” title.

These are all titles in this dataset:

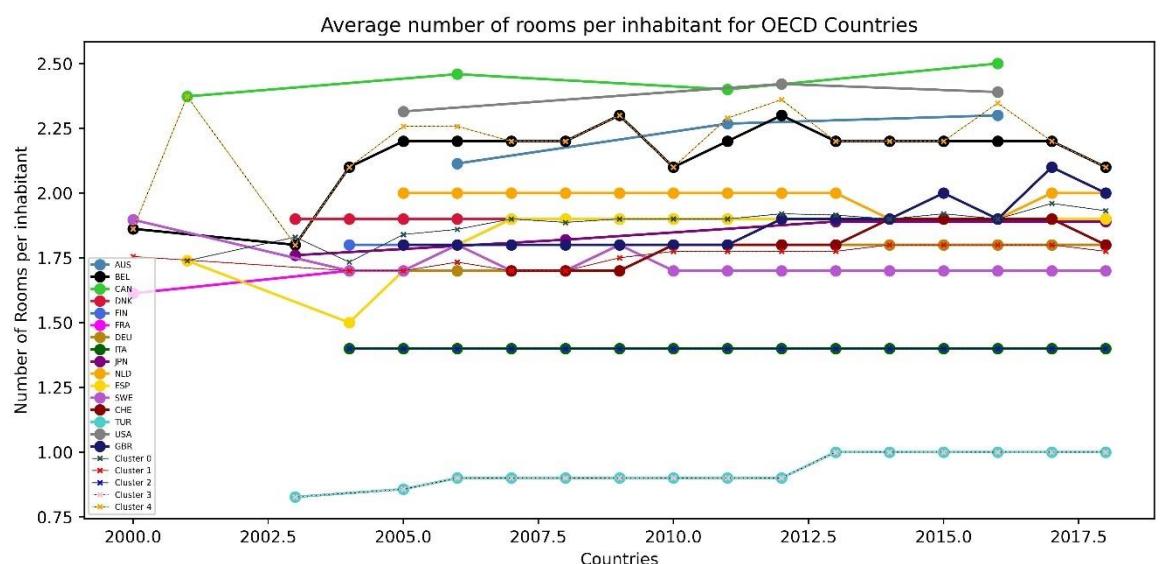
"TL", "Territory level and

Typology", "REG_ID", "Region", "VAR", "Indicator", "POS", "Position", "TIME", "Year",
"Unit Code", "Unit", "PowerCode Code", "PowerCode", "Reference Period
Code", "Reference Period", "Value", "Flag Codes", "Flags"

The number of rooms per person can help us understand how much privacy people have and how comfortable they live.

cluster	countries
Cluster 0	['DNK', 'FIN', 'JPN', 'NLD', 'ESP', 'GBR']
Cluster 1	['FRA', 'DEU', 'SWE', 'CHE']
Cluster 2	['ITA']
Cluster 3	['TUR']
Cluster 4	['AUS', 'BEL', 'CAN', 'USA']

We see Türkiye and Italy alone in the clusters.



- Türkiye is at the bottom of the chart. The reasons for the low number of rooms per person may be that many people stay in the same house or rooms due to the housing shortage in the country or the high rent and house prices.
- Italy has a fixed and lower rate than other countries. They are in a separate cluster by itself, as there is no similar data with themself.

In countries where house rents and house prices are high, the number of people living in the same house is high, as expected. So, the number of rooms per person is really low. The negative effects of poor housing conditions on children can continue throughout life, determine their future socioeconomic status, and affect adult wellbeing (Solari,2012). This factor, which also affects the future of children, is actually quite important. For this reason, children who grow up in a not very crowded house and have their own room can be expected to be more successful in the future.

5. Income and Wealth (Finance)

Although money does not buy happiness, it allows us to achieve higher living standards and a comfortable life. Money is also a way to get the best quality in many areas such as health, education, security. The higher economic wealth of the citizens is a key to the well-being of the country.

We can measure the economic situation of the country by GDP Growth Per Capita over the years.

We can see the annual income and expenditure balance of the government with the General Government Deficit.

We can observe the strength of the country's currency with Exchange Rates.

Even if a country's income is high, the distribution of this income to the public is also important. Income inequality in the society can lead to the polarization of the country.

All these factors can be informative in terms of understanding the economic situation of the country.

a. GDP per Capita Growth

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Productivity" section with "Growth in GDP per capita, productivity and ULC" title.

These are all titles in this dataset:

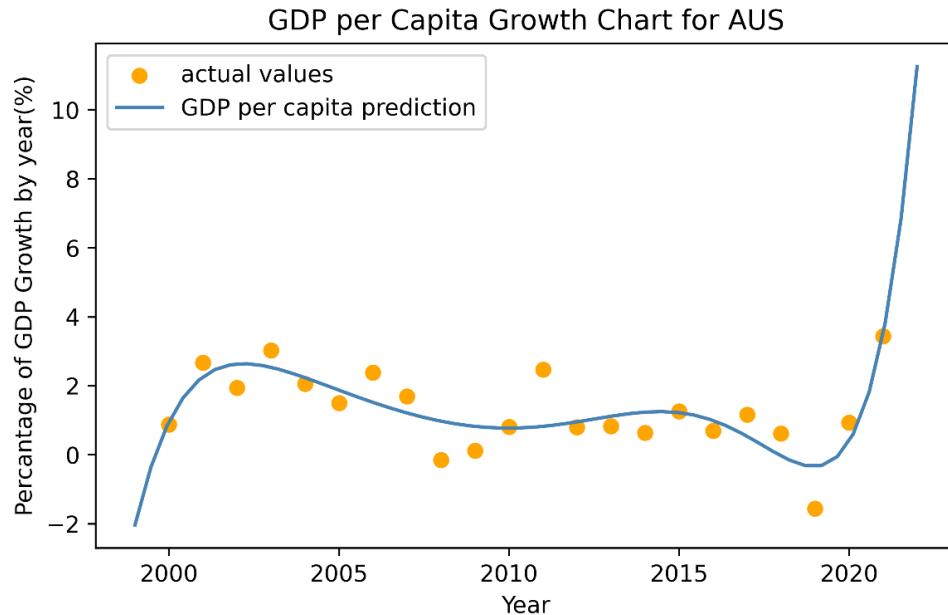
"LOCATION","Country","SUBJECT","Subject","MEASURE","Measure","TIME","Time","Unit Code","Unit","PowerCode Code","PowerCode","Reference Period Code","Reference Period","Value","Flag Codes","Flags"

Gross domestic product (GDP) is all the products and services produced and sold by a country. GDP per capita is calculated by dividing the GDP by the population of that country. The increases and decreases of GDP per Capita over the years help us to understand the changes in the economic situation of the country.

cluster	countries
Cluster 0	['AUS']
Cluster 1	['DNK', 'FIN', 'JPN', 'SWE', 'USA']
Cluster 2	['BEL', 'CAN', 'FRA', 'ITA', 'ESP', 'GBR']
Cluster 3	['DEU', 'NLD', 'CHE']
Cluster 4	['TUR']

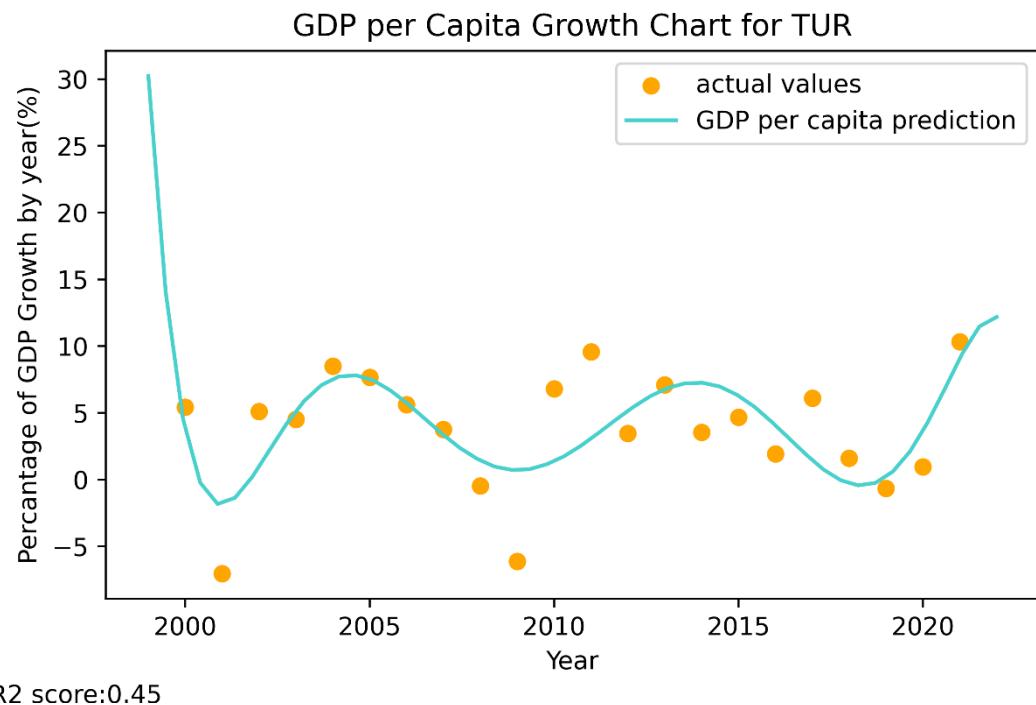
We see Australia and Türkiye in their clusters alone. Let's examine them:

- **Australia**



- Although we see a slightly decreasing trend in the growth of GDP per capita, we can say that it started to rise again after COVID-19. The R2 score may be informative, but not reliable with a value of 0.45.

- **Türkiye**



- We can say that Türkiye is an unstable country. That's why the R2 score is low, like 0.45.

It is important that a country's GDP is stable and growing. Because a stable economy is indispensable for a prosperous country. Looking at these graphs, it is possible to say that Turkey is a more unstable country than others. A high and stable GDP affects almost all the factors that affect the well-being of countries. In addition, the distribution of GDP within the population is also very important.

b. General Government Deficit

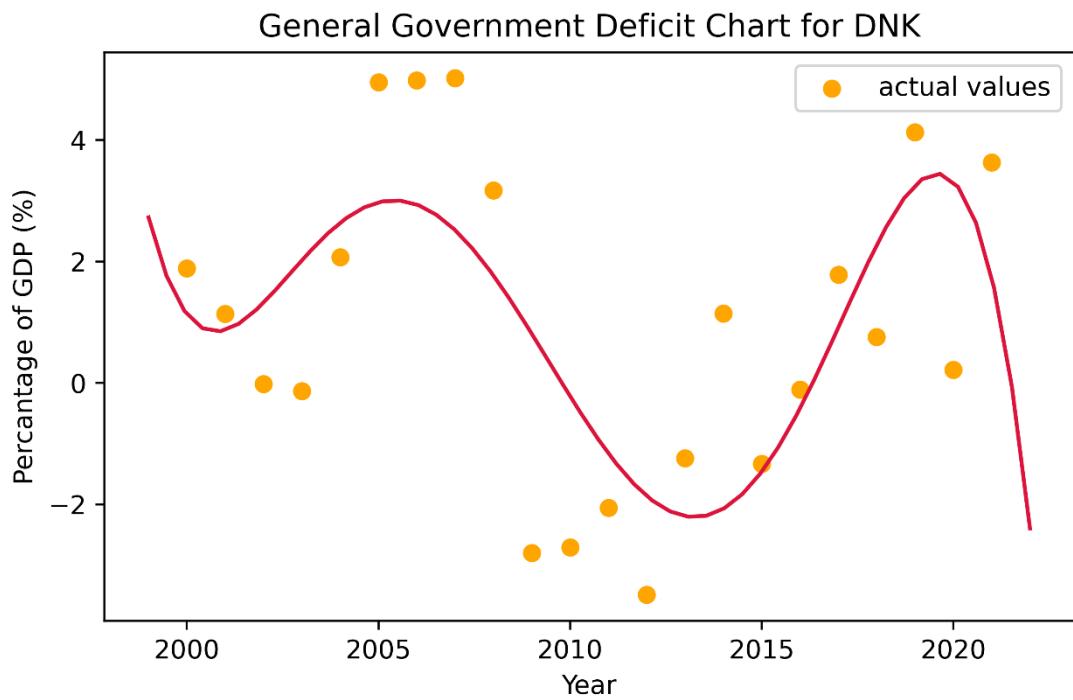
This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "National Accounts" section with "Government deficit/surplus, revenue, expenditure and main aggregates" title. These are all titles in this dataset:
 "LOCATION", "INDICATOR", "SUBJECT", "MEASURE", "FREQUENCY", "TIME", "Value", "Flag Codes"

General Government Deficit is the balance of income and expenditure of the country in that year. We can use it to understand whether that country is making a profit or a loss.

cluster	countries
Cluster 0	['DNK']
Cluster 1	['FIN', 'SWE']
Cluster 2	['DEU']
Cluster 3	['CHE']
Cluster 4	['AUS', 'BEL', 'CAN', 'FRA', 'ITA', 'JPN', 'NLD', 'ESP', 'TUR', 'USA', 'GBR']

We see Denmark, Germany, Switzerland alone in their clusters and Finland and Sweden in same cluster.

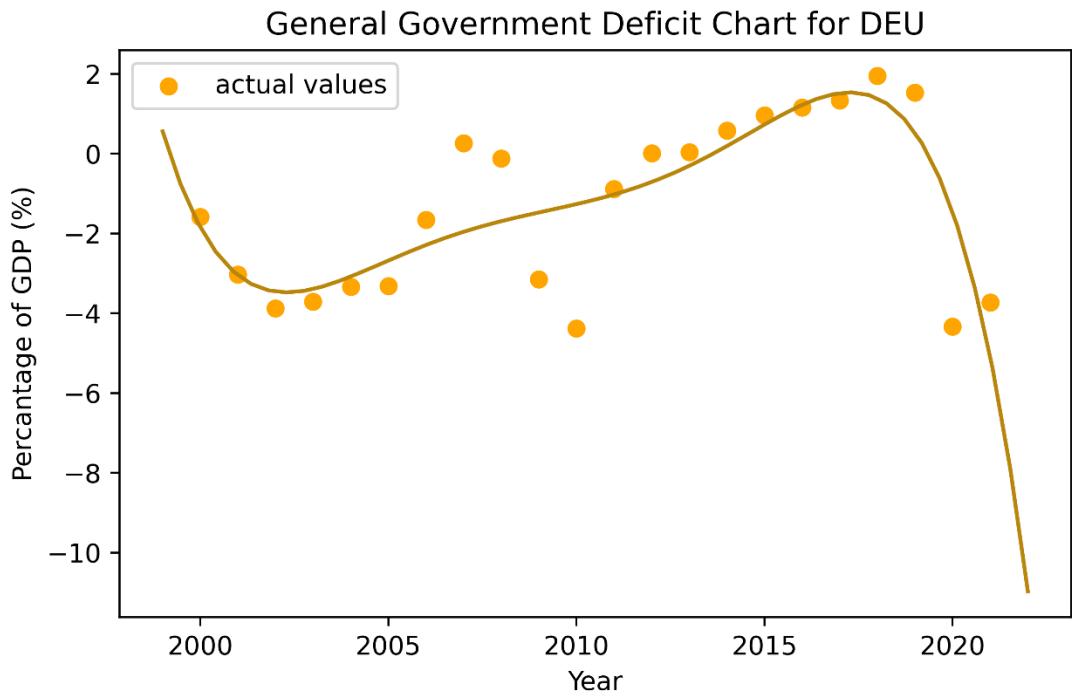
- **Denmark**



R2 score: 0.48

- We see an unstable country with increases and decreases. It is seen that even after years of profit, losses are incurred. Profit and loss ratio ranges from +4% to -3%.

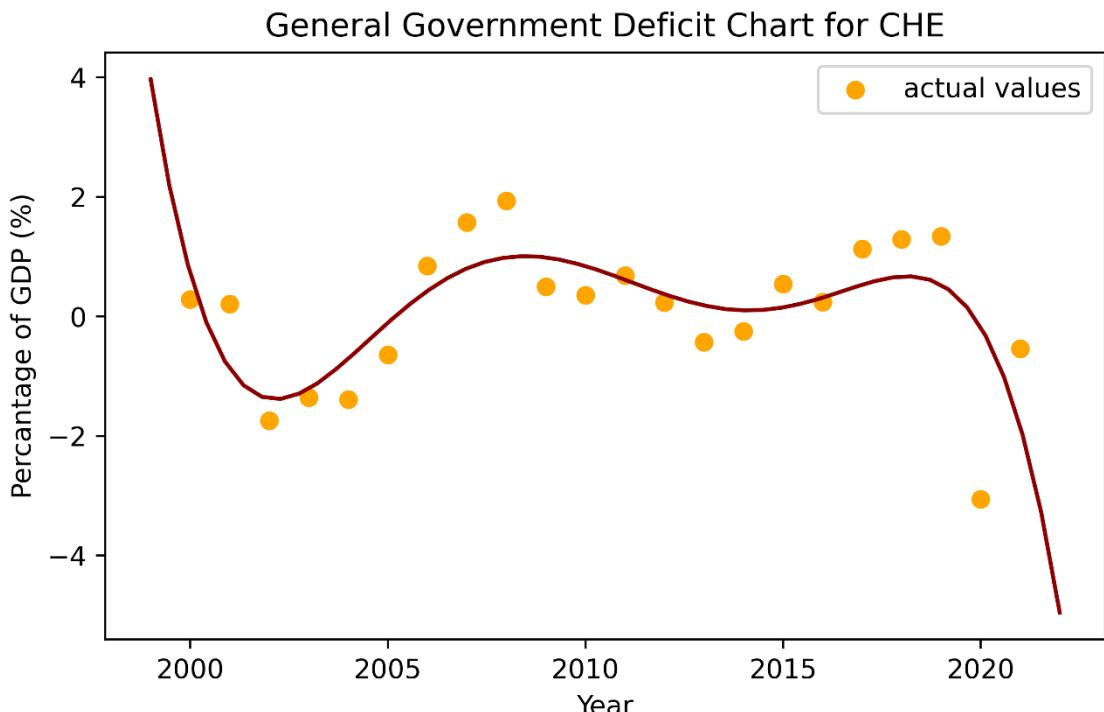
- **Germany**



R2 score: 0.66

- The country started with a 3% loss and recovered well. We're seeing a graph where the loss is decreasing over time, and even though the country profits in 2019-2020, we're seeing a sharp decline post covid-19.

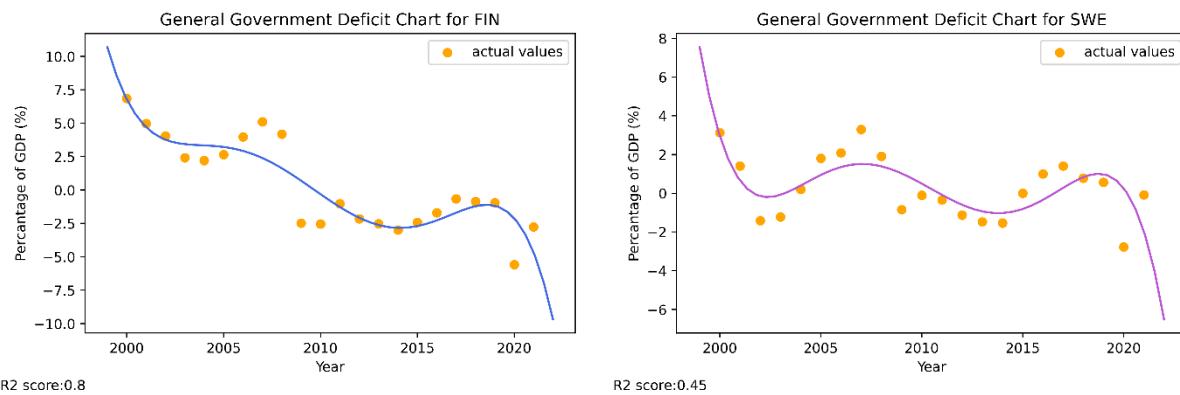
- **Switzerland**



R2 score: 0.45

- We see a graph that varies between -2% and +2%. Although we see a very unchanged graphic in general, it is possible to say that there is a noticeable decrease after COVID 19.

- **Finland and Sweden**



- Finland is a country in a downward trend due to high profit rates. Sweden also appears to be in a slightly downtrend. Although the trends of these two countries are not very similar, it should not be surprising that these two countries are in the same cluster as the profit and loss values after 2010 are very similar.

The fact that countries do not have a GDP deficit is important in terms of ensuring the stability of that country's economy. According to research by Hassan et al., high unemployment, interest rate and inflation rate have a negative effect on the country's GDP deficit (2014). These are just a few of the factors that matter to the GDP deficit. It is essential to follow a safe economic policy by paying attention to these factors, to ensure the balance of the welfare of the country.

c. Income Inequality

This dataset is retrieved from "OECD Data" as ".csv" file. This dataset can be reached from <https://data.oecd.org/inequality/income-inequality.htm> with "Income inequality" title.

These are all titles in this dataset:

"LOCATION", "INDICATOR", "SUBJECT", "MEASURE", "FREQUENCY", "TIME", "Value", "Flag Codes"

Many tools can be used to measure Income Inequality. Some of these are those:

Gini coefficient: ranges between 0 in the case of perfect equality and 1 in the case of perfect inequality.

S80/S20: The ratio of the average income of the 20% richest to the 20% poorest.

P90/P10: The ratio of the upper bound value of the ninth decile (i.e., the 10% of people with highest income) to that of the first decile.

P90/P50: The upper bound value of the ninth decile to the median income.

P50/P10 of median income to the upper bound value of the first decile.

The Palma ratio is the share of all income received by the 10% people with highest disposable income divided by the share of all income received by the 40% people with the lowest disposable income.

Let's see clusters for these inequalities:

Clusters for PALMA:

cluster	countries
Cluster 0	['CAN', 'FRA', 'DEU', 'NLD', 'CHE']
Cluster 1	['BEL', 'DNK', 'FIN', 'SWE']
Cluster 2	['AUS', 'ITA', 'JPN', 'ESP']
Cluster 3	['TUR', 'USA', 'GBR']

Clusters for P90/P50:

cluster	countries
Cluster 0	['CAN', 'FRA', 'DEU', 'NLD', 'CHE']
Cluster 1	['AUS', 'ITA', 'JPN', 'ESP', 'GBR']
Cluster 2	['BEL', 'DNK', 'FIN', 'SWE']
Cluster 3	['TUR', 'USA']

Clusters for P90/P10:

cluster	countries
Cluster 0	['JPN', 'ESP', 'TUR', 'USA']
Cluster 1	['BEL', 'DNK', 'FIN', 'FRA', 'NLD', 'SWE']
Cluster 2	['CAN', 'DEU', 'CHE']
Cluster 3	['AUS', 'ITA', 'GBR']

Clusters for P50/P10:

cluster	countries
Cluster 0	['BEL', 'DNK', 'FIN', 'FRA', 'NLD', 'SWE', 'CHE']
Cluster 1	['AUS', 'CAN', 'DEU', 'GBR']
Cluster 2	['ITA', 'ESP', 'TUR']
Cluster 3	['JPN', 'USA']

Clusters for GINI:

cluster	countries
Cluster 0	['BEL', 'DNK', 'FIN', 'SWE']
Cluster 1	['CAN', 'FRA', 'DEU', 'NLD', 'CHE']
Cluster 2	['AUS', 'ITA', 'JPN', 'ESP', 'GBR']
Cluster 3	['TUR', 'USA']

Clusters for S80/S20:

cluster	countries
Cluster 0	['BEL', 'DNK', 'FIN', 'SWE']
Cluster 1	['CAN', 'FRA', 'DEU', 'NLD', 'CHE']
Cluster 2	['AUS', 'ITA', 'JPN', 'ESP', 'GBR']
Cluster 3	['TUR', 'USA']

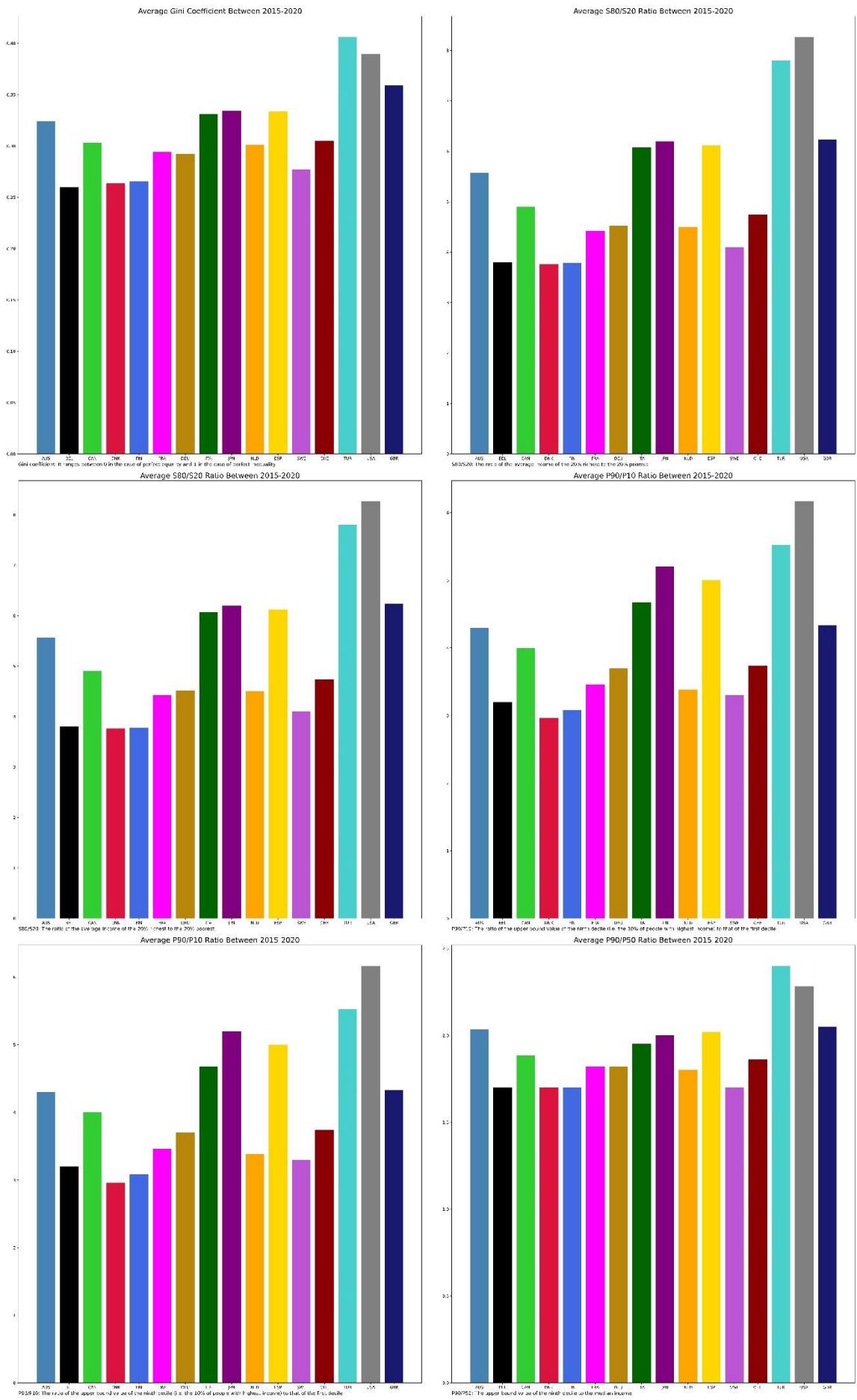
When we look overall these clusters, we can say these:

Türkiye and USA are mostly in same clusters.

Belgium, Denmark, Finland, and Sweden are mostly in same clusters.

Canada, Germany, and Netherlands are mostly in same clusters.

Let's see chart for details:



- Income inequality in Türkiye and USA is very high.
- Income inequality in Belgium, Denmark, Finland, and Sweden is very low.

High income inequality indicates that there are very rich and very poor people in the society in that country. The rich own most of the property in the country. In these countries, the middle-income group is low.

The low-income inequality indicates that most of the people in that country are from the middle class and there is not much difference between people in terms of income. We can say that there is a more equitable distribution in these countries.

According to Ravallion's study in 2014, although income inequality throughout the world decreased until the 2000s, we see that this inequality increased somewhat in the 2000s. In addition, Ravallion argues that lawmakers can provide the balance against poverty. Ravallion's view is quite accurate. Because without government intervention, inequality will continue to increase. The government should provide more assistance to the poor with a social understanding and ensure equal opportunities for children and poor by taxing the wealthy more.

d. Exchange Rates

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "National Accounts" section with "PPPs and exchange rates" title.

These are all titles in this dataset:

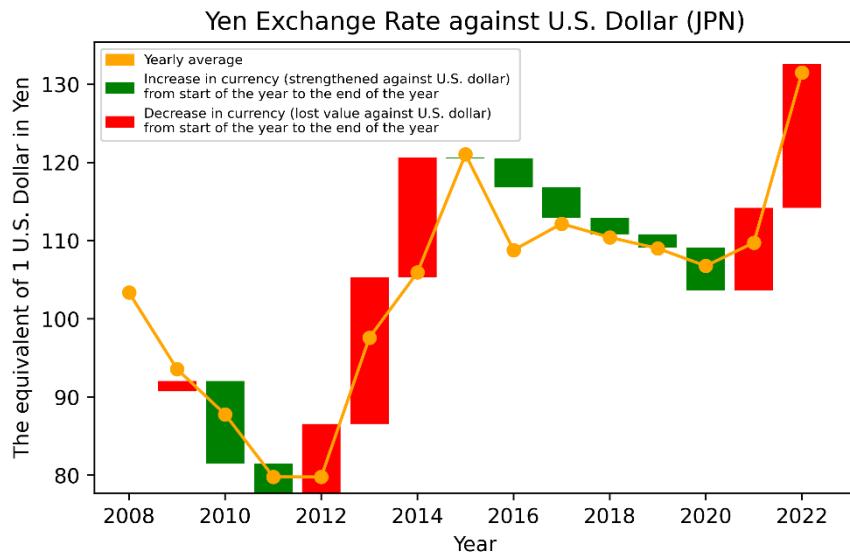
"LOCATION", "Country", "TRANSACT", "Transaction", "MEASURE", "Measure", "TIME", "Year", "Unit Code", "Unit", "PowerCode Code", "PowerCode", "Reference Period Code", "Reference Period", "Value", "Flag Codes", "Flags"

The exchange rate is the strength of that country's currency against other currencies. In this dataset, we will examine the changes in the currencies of the countries against the dollar over the years, ignoring the depreciation of the dollar with inflation.

cluster	countries
Cluster 0	['AUS', 'CAN', 'SWE']
Cluster 1	['BEL', 'DNK', 'FIN', 'FRA', 'DEU', 'ITA', 'NLD', 'ESP', 'GBR']
Cluster 2	['JPN']
Cluster 3	['TUR']
Cluster 4	['CHE', 'USA']

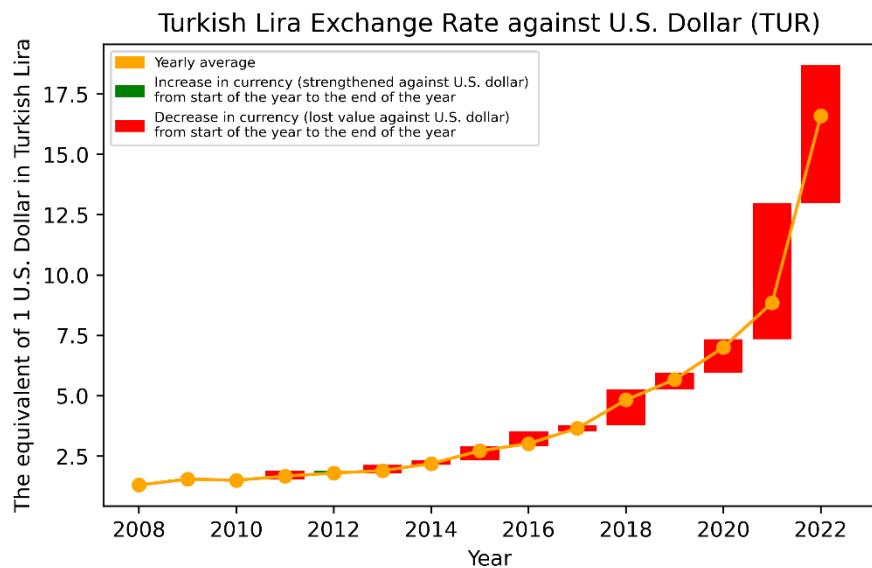
Cluster 1 are European countries that use the Euro as their currency.
We see Japan and Türkiye as alone in their clusters.

Yen (Currency of Japan)



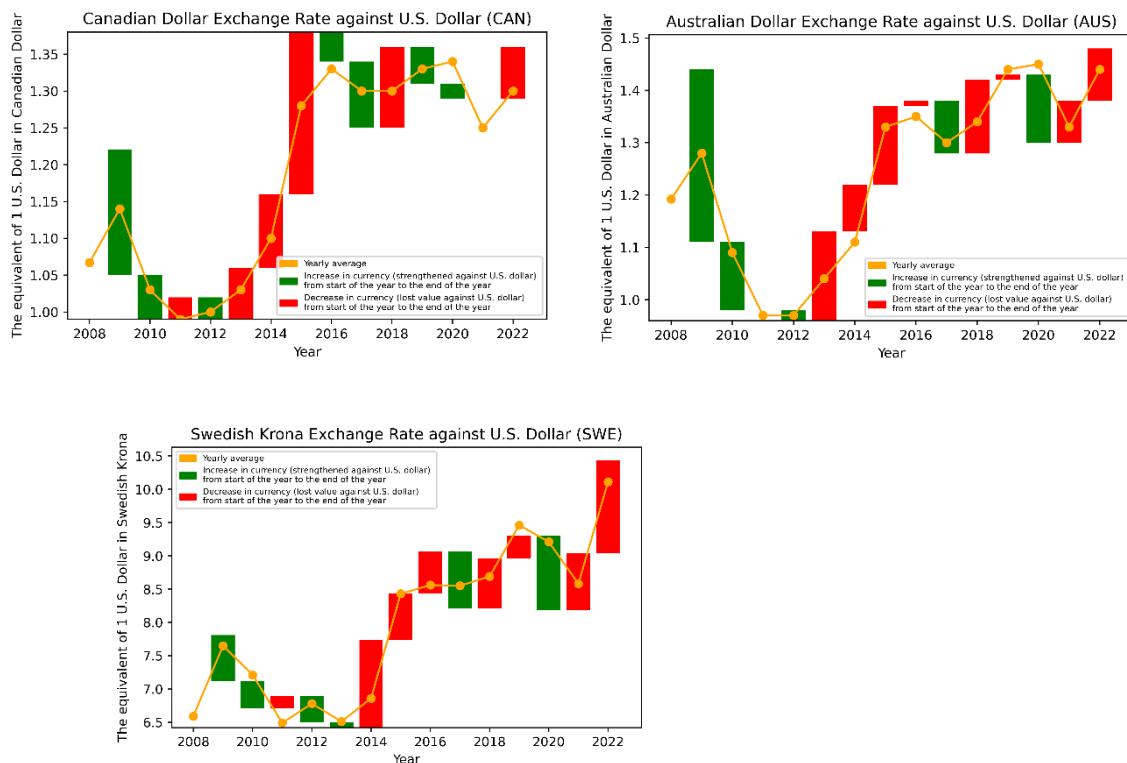
- The yen appreciated against the dollar until 2012. Afterwards, we see that the Yen depreciated until 2015, and although it tried to recover between 2015-2020, it lost value again afterwards.

Turkish Lira (Currency of Türkiye)



- We see that the Turkish lira has been losing value against the dollar since 2008. Especially the depreciation after 2018 and 2020 is very remarkable. In 2021, the Turkish lira depreciated by half against the dollar. This depreciation seems to have continued in 2021 as well.

Canadian Dollar, Australian Dollar, and Swedish Krona (Cluster 0)



- When we look at the yearly average graph of the currencies of these 3 countries against the dollar, we see 3 very similar graphs, although the graphs are not exactly the same.

The exchange rate is the strength of one country's currency against the other country's currency. In other words, a strong and exporting economy means a strong currency. For the country's currency to appreciate, it must bring money into the country. This is possible with export. In other words, a country that exports more than imports has earned money from foreign trade. For this reason, it is imperative that the country produces and has a strong economy. The Swiss franc has been in almost constant value for years against the dollar. The government follows a policy that can be an example for other countries by managing the country policy of money perfectly. If a country imports more than it exports, it has a fragile economy. If problems arise in the economy of the country, this will have very devastating consequences for the economy. The best example of this is the depreciation of the Turkish currency. It is depreciating exponentially against other countries. The government should make more exports by turning to the right policies.

6. Knowledge and Skills (Education)

Education is very important for people. It should not be expected that every individual in a society has received a very high level of education. Because this situation can create problems such as qualification mismatch and overqualification. Likewise, the high number of uneducated individuals in a society can lead to problems such as security, unemployment and underqualification. That's why education in the right way and in the right field is very important.

a. PISA

3 different datasets used in this field:

All these datasets are retrieved from “OECD Data” as ".csv" file.

Math PISA scores are retrieved from <https://data.oecd.org/pisa/mathematics-performance-pisa.htm> with “Mathematics performance (PISA)” title.

Read PISA scores are retrieved from <https://data.oecd.org/pisa/reading-performance-pisa.htm> with “Reading performance (PISA)” title.

Science PISA scores are retrieved from <https://data.oecd.org/pisa/science-performance-pisa.htm> with “Science performance (PISA)” title.

Average scores are calculated by average of these scores.

These are all titles in this dataset:

“LOCATION”, “INDICATOR”, “SUBJECT”, “MEASURE”, “FREQUENCY”, “TIME”,
“Value”, “Flag Codes”

The Program for International Student Assessment (PISA) is an assessment to measure the Reading, Mathematics and Science knowledge of 15-year-olds in OECD countries.

This is a good metric to evaluate the education provided by this country and to compare it with other countries.

Let's see Clusters in these fields:

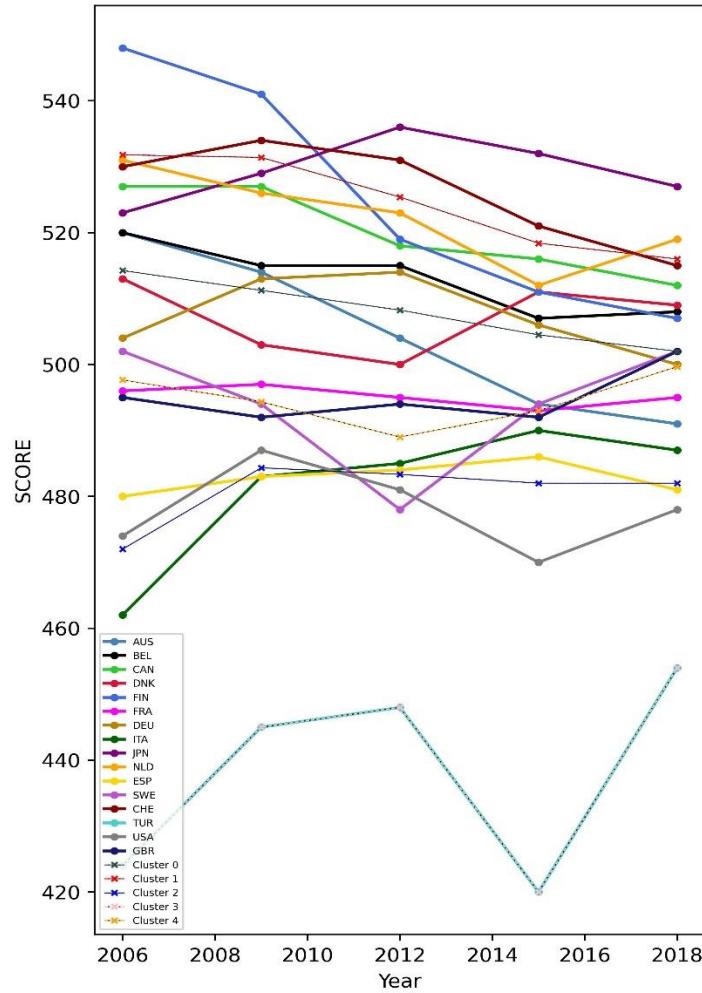
Math

Mathematics is an indispensable part of education life. Because mathematics is used a lot in daily life. Apart from this, many skills such as making logical and correct decisions develop thanks to mathematics in education. So, Mathematics is very important.

cluster	countries
Cluster 0	['AUS', 'BEL', 'DNK', 'DEU']
Cluster 1	['CAN', 'FIN', 'JPN', 'NLD', 'CHE']
Cluster 2	['ITA', 'ESP', 'USA']
Cluster 3	['TUR']
Cluster 4	['FRA', 'SWE', 'GBR']

Türkiye is an outlier in these clusters. Other countries are clustered among themselves.

Programme for International Student Assessment
PISA MATH RESULTS



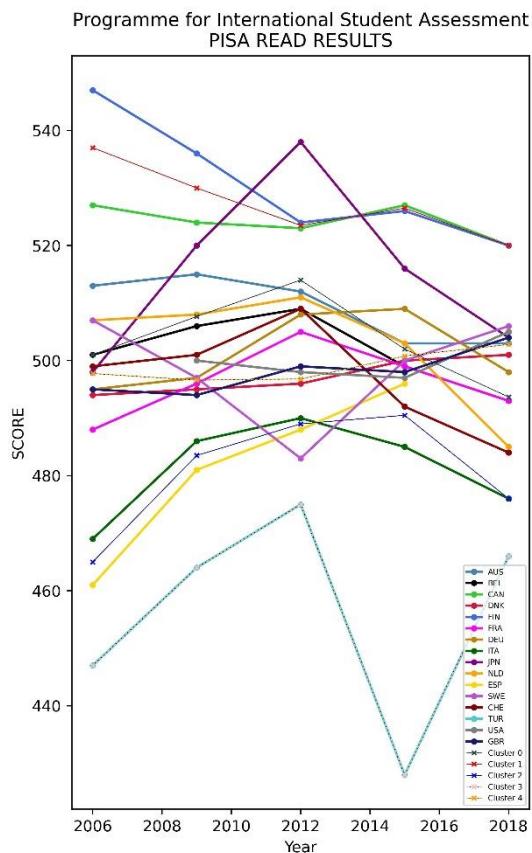
- As we see in clusters, Türkiye is an outlier in this graph. Math scores for Türkiye is really low. Türkiye should improve its education in maths much more.

Read

Language is very important for communication. We communicate with language. When we communicate, we exchange information. The ways of uploading information are by speaking or writing. The knowledge gained by reading is more permanent. Therefore, reading is very important for acquiring knowledge. Understanding what you read is also essential for obtaining the correct information. In societies that do not understand what they read, many problems are encountered, starting with communication.

cluster	countries
Cluster 0	['AUS', 'BEL', 'FRA', 'JPN', 'NLD', 'CHE']
Cluster 1	['CAN', 'FIN']
Cluster 2	['ITA', 'ESP']
Cluster 3	['TUR']
Cluster 4	['DNK', 'DEU', 'SWE', 'USA', 'GBR']

Türkiye seems to be outlier in this field too.



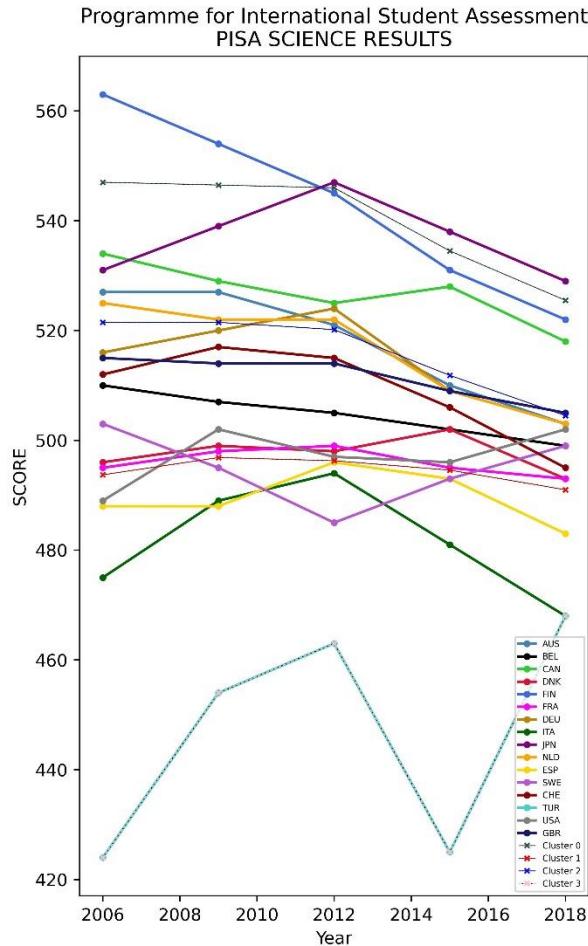
- Although Turkey has increased its score towards 2012, we see that it got a lower score in 2015 than in 2006. We can clearly say that people in Turkey do not understand what they read.
- While Canada and Finland (Cluster 1) are at the top, it is possible to say that their scores also dropped a little.
- Although Italy and Spain got close scores until 2012, we see that the difference has changed after 2012.

Science

Science is essential to the development of civilizations. Today, we do a lot of things thanks to science. Science should be taught to children at an early age and children should learn science with love. A country that is not good at science cannot develop.

cluster	countries
Cluster 0	['FIN', 'JPN']
Cluster 1	['BEL', 'DNK', 'FRA', 'ITA', 'ESP', 'SWE', 'USA']
Cluster 2	['AUS', 'CAN', 'DEU', 'NLD', 'CHE', 'GBR']
Cluster 3	['TUR']

Türkiye is an outlier in these clusters. Another different cluster is Cluster 0.



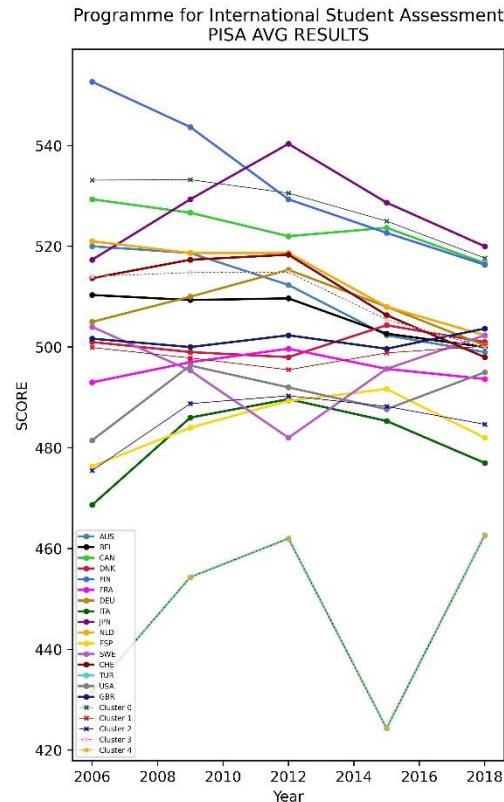
- We see that Turkey is at the bottom in the field of science, as in other fields.
- While Finland and Japan had different scores at the beginning, we see that their scores are quite close after 2012.

Average

This data is average of Math, Read, Science scores.

cluster	countries
Cluster 0	['CAN', 'FIN', 'JPN']
Cluster 1	['DNK', 'FRA', 'SWE', 'GBR']
Cluster 2	['ITA', 'ESP', 'USA']
Cluster 3	['AUS', 'BEL', 'DEU', 'NLD', 'CHE']
Cluster 4	['TUR']

Türkiye was an outlier in all these Clusters. It is not surprising they are again outlier in Average results.



- Although Canada, Finland and Japan are at the top in separate fields, they are all together at the top when looking at the average score. Because they were in the middle even in the field where they were not at the top.
- Since Turkey is in the last place in all fields, it is not surprising that it is at the bottom when the average score is considered. Turkey needs to progress in the field of education by examining the education system of these top countries. Because it seems that students do not receive a very correct education at school.

We see which countries' children are more successful by looking at the PISA results, but this is not a complete reflection of the education system. Many factors, such as the children's way of life, nutrition, social life, and character, influence these results. However, these results are still important in terms of providing an idea. When we look at these results, we see that children in Turkey perform at much lower levels compared to other countries. Although many factors play a role in this, we cannot say that the education system is correct. While there are many factors contributing to the success of children growing up in countries like Finland and Canada, they definitely have successful education systems. In addition, Bulle says in his research that Finland's education system is taken as an example by other countries and that the school policies in Western countries are based on comparison, which is the source of the failure (2011). It is highly misguided to start comparing and categorizing children as soon as they begin their educational journey.

b. Student per Instructor

This dataset is retrieved from "OECD Data" as ".csv" file. This dataset can be reached from <https://data.oecd.org/teachers/students-per-teaching-staff.htm> with "Students per teaching staff" title.

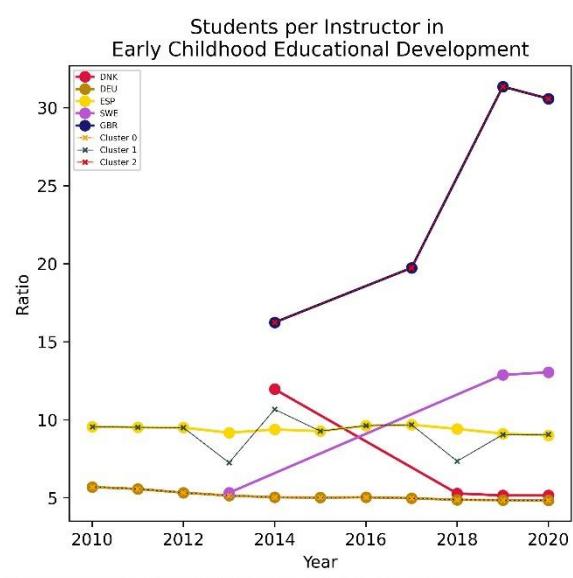
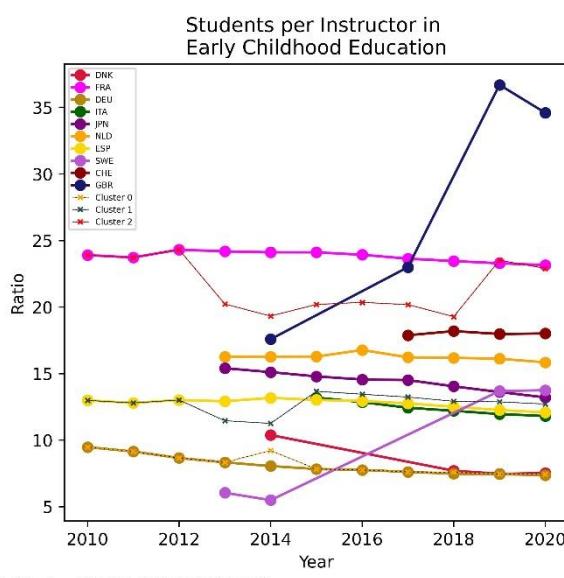
These are all titles in this dataset:

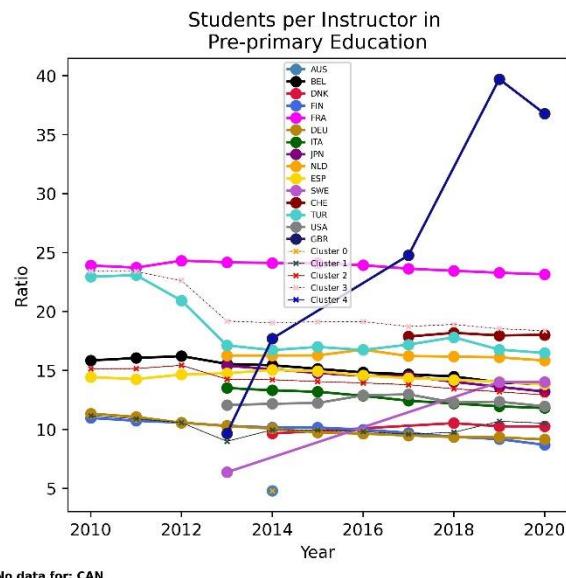
"COUNTRY", "Country", "INDICATOR", "Indicator", "EDUCATION_LEVEL", "Education level", "REF_SECTOR", "Reference sector", "YEAR", "Year", "Value", "Flag Codes", "Flags"

Student per instructor is a good metric in education. Because if an instructor has too many students, then instructor cannot devote a lot of time to each student. Specially,

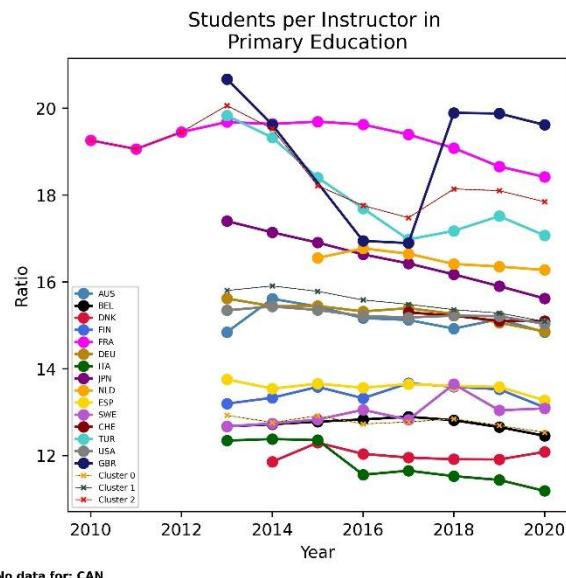
This is important especially at lower levels such as primary school.

These are the graphs with clusters for every education level and every education types:

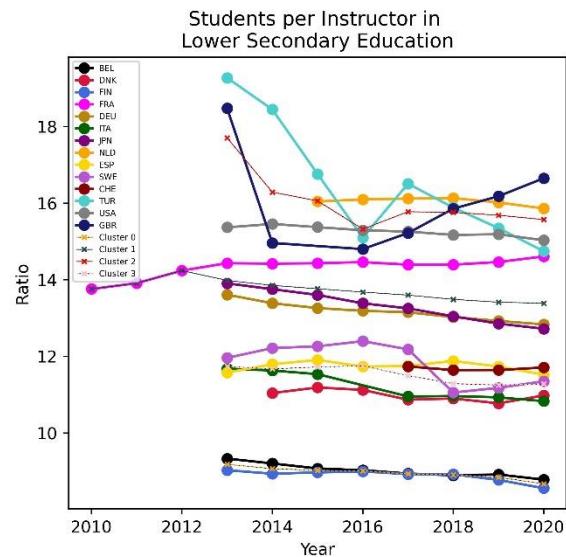




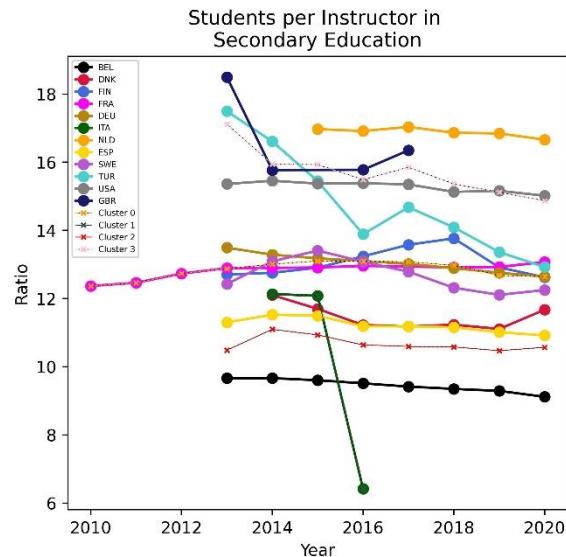
No data for: CAN



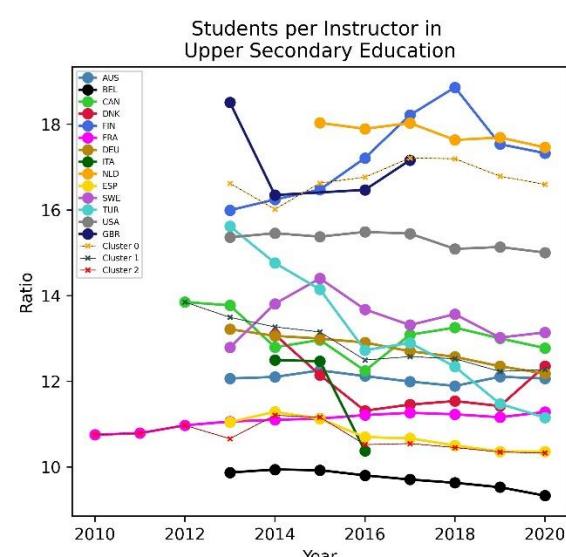
No data for: CAN



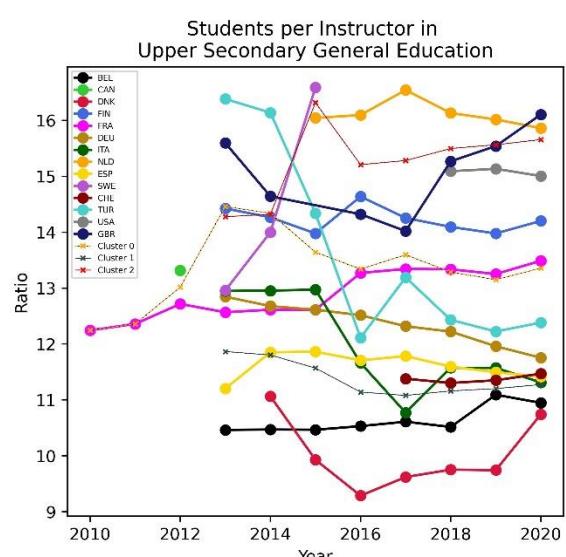
No data for: AUS, CAN



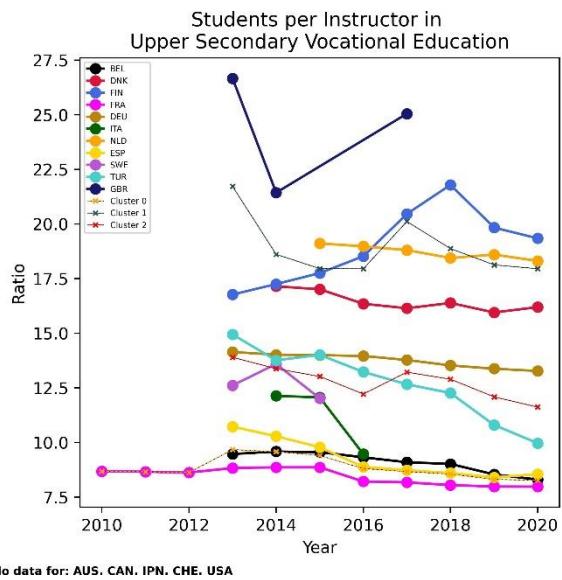
No data for: AUS, CAN, IPN, CHE



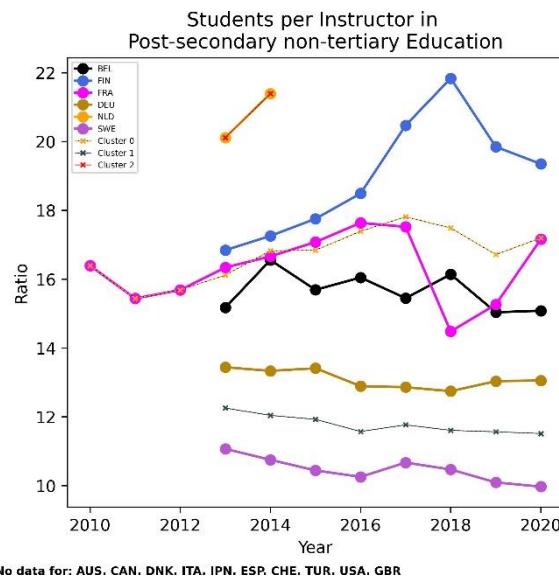
No data for: IPN, CHE



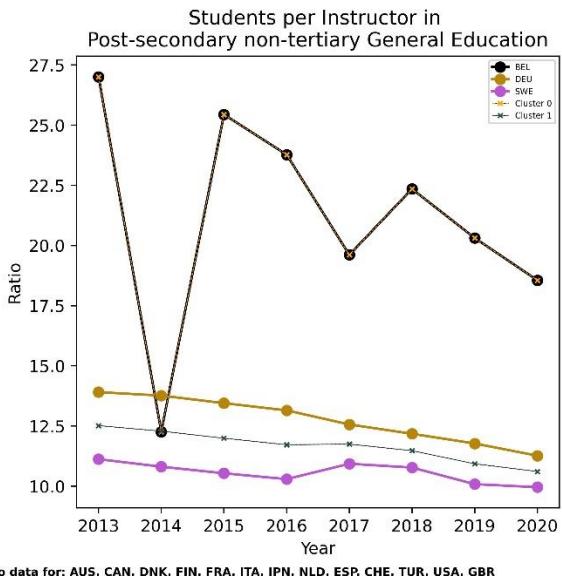
No data for: AUS, CAN



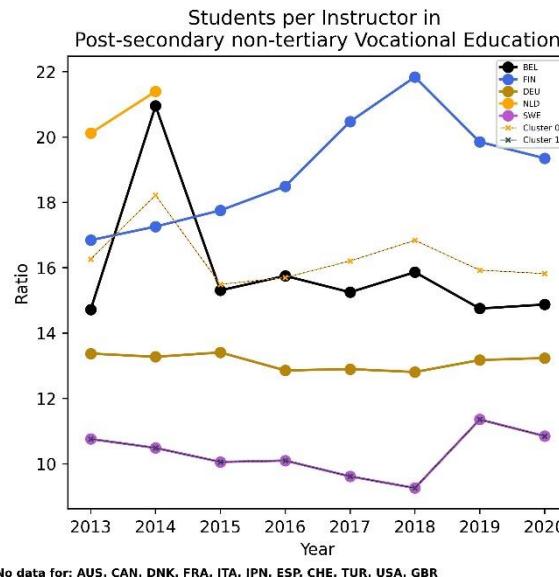
No data for: AUS, CAN, IPN, CHE, USA



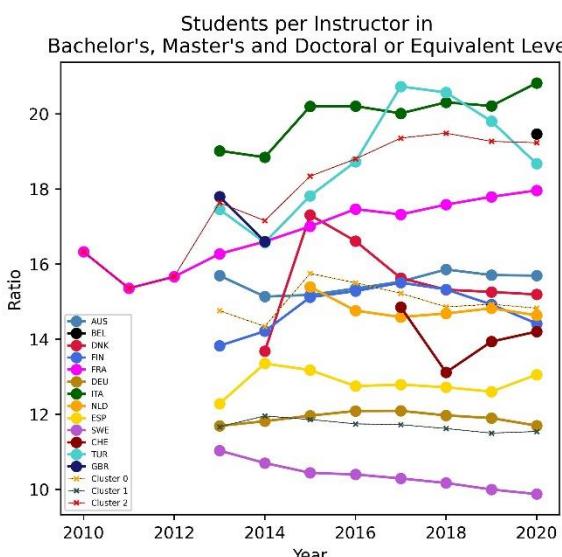
No data for: AUS, CAN, DNK, ITA, IPN, ESP, CHE, TUR, USA, GBR



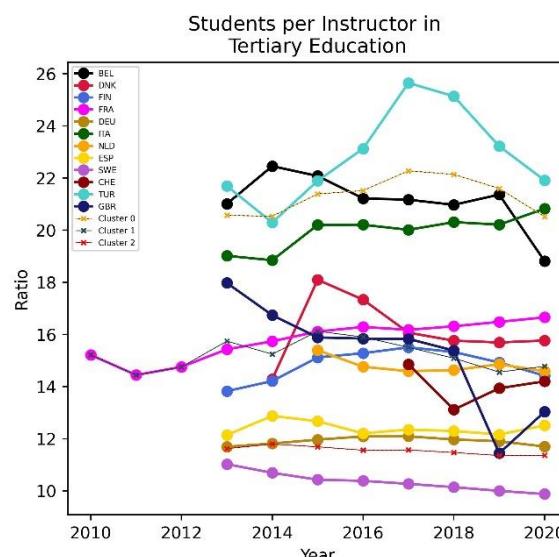
No data for: AUS, CAN, DNK, FIN, FRA, ITA, IPN, NLD, ESP, CHE, TUR, USA, GBR



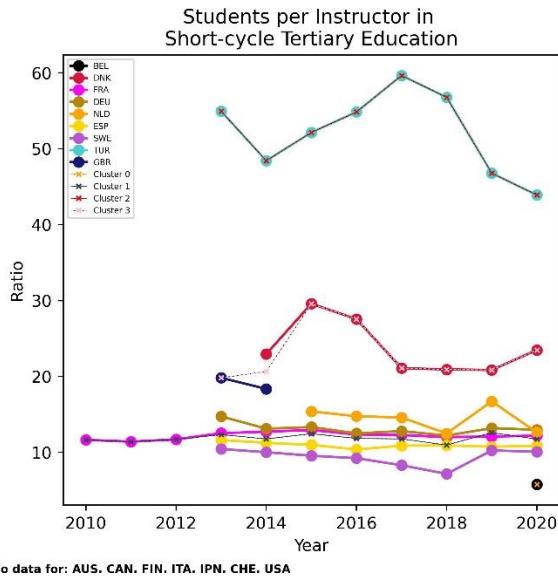
No data for: AUS, CAN, DNK, FRA, ITA, IPN, ESP, CHE, TUR, USA, GBR



No data for: CAN, IPN, USA



No data for: AUS, CAN, IPN, USA



In general, the outstanding data are:

- Türkiye ranks high in the number of instructors per capita. So, an instructor deals with too many students. This may be the reason why they fail in PISA and students do not get a good education.
- The number of students per instructor in Sweden is quite low. So, students get a good education and get good results in PISA.
- In early childhood education, the number of students per instructor in Germany is quite low.

The number of instructors per student is a very important factor, although the data we have is very diverse due to the differences in terms. However, it is sufficient for us to comment in general. As the number of students, a teacher is responsible for increases, the quality of education they receive decreases, leading to an expected outcome of failure. The research conducted by Koc et al. has revealed that an increase in the instructor per student ratio negatively affects success. Despite being costly, they have stated that hiring more teachers would be the right move as it would reduce the workload of teachers and enhance students' performance (2015). This recommendation is quite correct. Because education is a subject where all kinds of expenditures must be made for the order and future of society. It is not possible for an uneducated society to get better in the future.

c. Skill Mismatch

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Labour" section with "Mismatch by country" title. These are all titles in this dataset:

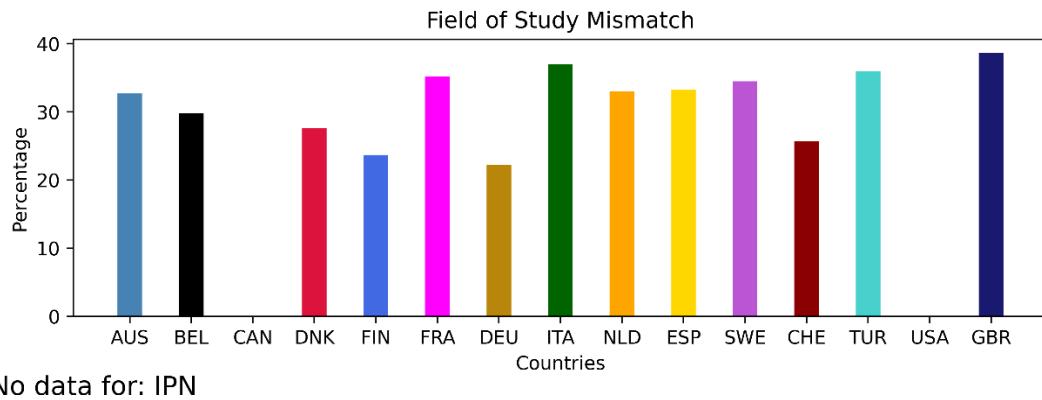
"LOCATION", "Country", "MISMATCH", "Mismatch", "Value", "Flag Codes", "Flags"

Skill Mismatch is the situation where the skills of the employee and the needs of the job are incompatible. This may be due to:

1. Field of Study Mismatch

This type of mismatch means that the employee is working in a field different from his/her graduation. This may be because he has not been able to find a job in his field or because he has had a better opportunity.

cluster	countries
Cluster 0	['DNK', 'FIN', 'DEU', 'CHE']
Cluster 1	['AUS', 'BEL', 'NLD', 'ESP']
Cluster 2	['FRA', 'ITA', 'SWE', 'TUR', 'GBR']
Cluster 3	['CAN', 'USA']



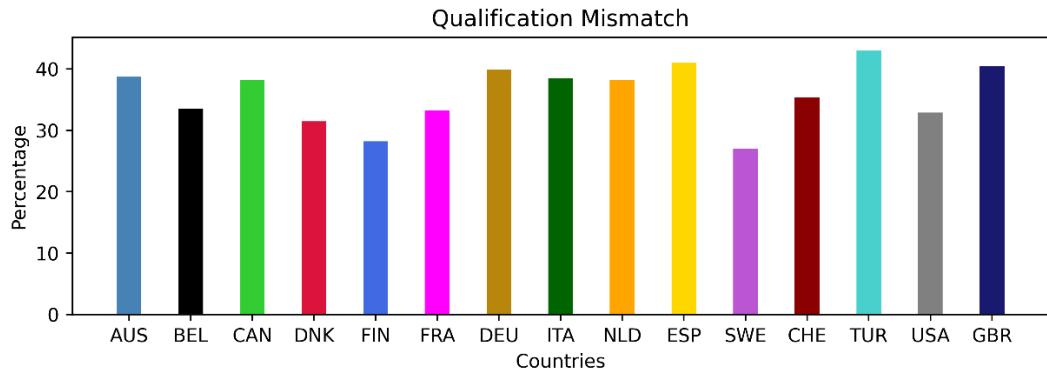
- This chart is prepared based on the data OECD provided. USA and Canada have 0% “Field of Study Mismatch” on the data. We assume that this was due to an error.
- France, Italy, Sweden, Türkiye, and Great Britain (Cluster 2) has similar percent on this data, and they are on the top of the chart.

2. Qualification Mismatch

This type of incompatibility occurs when the employee cannot find a job in his field for a long time and works in a job independent of his field. Main reasons for this mismatch are Underqualification and Overqualification. We will see these closer.

Let's see data about Qualification Mismatch:

cluster	countries
Cluster 0	['AUS', 'CAN', 'ITA', 'NLD']
Cluster 1	['BEL', 'DNK', 'FRA', 'CHE', 'USA']
Cluster 2	['FIN', 'SWE']
Cluster 3	['DEU', 'ESP', 'TUR', 'GBR']



No data for: IPN

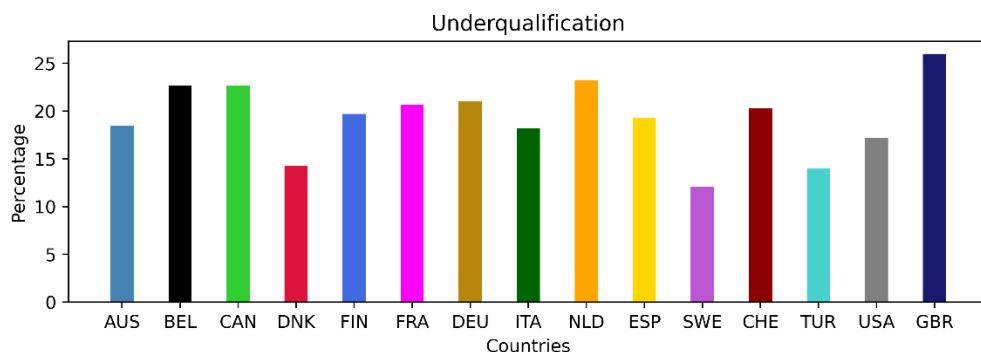
- Qualification Mismatch is 40 percent in Türkiye. But countries such as Spain, Germany and Great Britain are close to Türkiye.
- Finland and Sweden have a good educational system. The main reason for the low Competency Mismatch is that students should be trained according to their field of interest.

Let's see 2 Main reasons for this Qualification Mismatch:

a. Underqualification

Underqualification is the proportion of people who are insufficient in their field of work. The reason for this deficiency is insufficient education. The education here should not be considered as the education received only in schools. It is also possible for a master to train his apprentice to become a qualified worker.

cluster	countries
Cluster 0	['BEL', 'CAN', 'NLD', 'GBR']
Cluster 1	['AUS', 'ITA', 'USA']
Cluster 2	['FIN', 'FRA', 'DEU', 'ESP', 'CHE']
Cluster 3	['DNK', 'SWE', 'TUR']



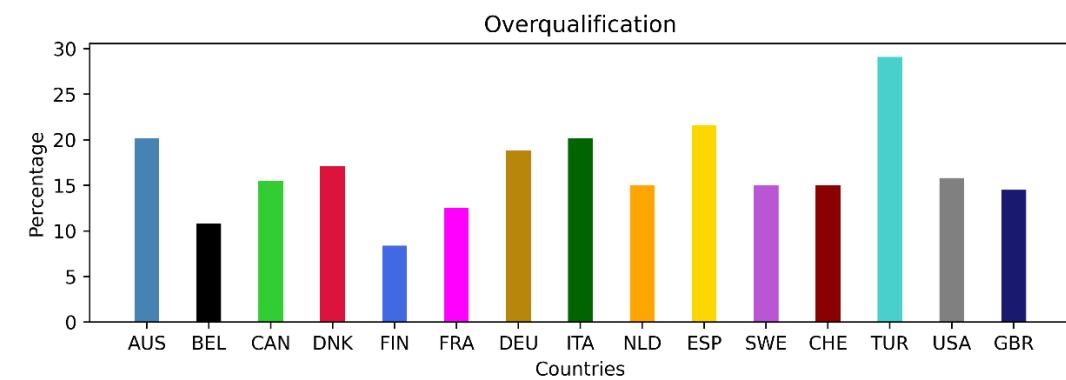
No data for: IPN

- When we look at the above chart and clusters, the number of insufficient people in their field is quite high in Belgium, Canada, Netherlands, and Great Britain.
- At the bottom of the chart are Denmark and Sweden, which stand out with their education. The fact that Turkey is located together with these two countries shows that the number of underqualified people is sufficient.

b. Overqualification

Overqualification is the proportion of people who are overly competent in their field of study. This may be because there are too many university graduates and less educated people in active positions.

cluster	countries
Cluster 0	['BEL', 'FIN', 'FRA']
Cluster 1	['AUS', 'DEU', 'ITA', 'ESP']
Cluster 2	['CAN', 'DNK', 'NLD', 'SWE', 'CHE', 'USA', 'GBR']
Cluster 3	['TUR']



No data for: IPN

- When we look at the table, Turkey is at the top of the overqualified rate by itself and by far. This may be because there are too many university graduates in Turkey. Although this seems like a good thing, less educated people are needed in every country. Countries should train people with fewer but better-quality education, instead of people who have received too much and less quality education. Another reason for this high rate may be that less educated individuals are in positions where educated individuals are required to work.
- Finland, which is successful in education, and Belgium and France, which are relatively successful, rank lowest in the proportion of overqualified people.
- The number of overqualified people is average even in Sweden, which is quite successful in education.

The government should educate the people according to the demand in the sector. According to Manacorda, if the government does not keep the supply-demand balance for jobs, there will be an unemployment problem in the country (2010). If there are more individuals trained for a profession than the demand for that profession, those individuals will become unemployed, and the unemployment rate will increase. On the other hand, if there are fewer individuals trained for a profession than the demand, the gap in that sector will continue to grow and employers in that field will struggle to find suitable employees and unemployment will rise due to an oversupply of workers in other job fields. The balance of supply and demand is very important for the unemployment rate to decrease. Therefore, it is essential to educate young people to be able to work in the right fields.

d. Education Over Job Quality

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Labour" section with "Job quality by skill" title.

These are all titles in this dataset:

"LOCATION", "Country", "MEA", "Overall measure", "VAR", "Components", "POP", "Age", "SEX", "Sex", "EDU", "Education", "TIME", "Time", "Unit Code", "Unit", "PowerCode Code", "PowerCode", "Reference Period Code", "Reference Period", "Value", "Flag Codes", "Flags"

The quality of the work people do is also very important. Every well-educated individual wants to work in a job where he can earn a good income. We will examine what employees have achieved by education and skill level in this dataset.

Clusters for Low Skilled:

cluster	countries
Cluster 0	['BEL', 'DNK', 'NLD', 'CHE']
Cluster 1	['AUS', 'FIN', 'FRA', 'DEU', 'ITA', 'SWE']
Cluster 2	['CAN', 'JPN', 'ESP', 'USA', 'GBR']
Cluster 3	['TUR']

Clusters for Medium Skilled:

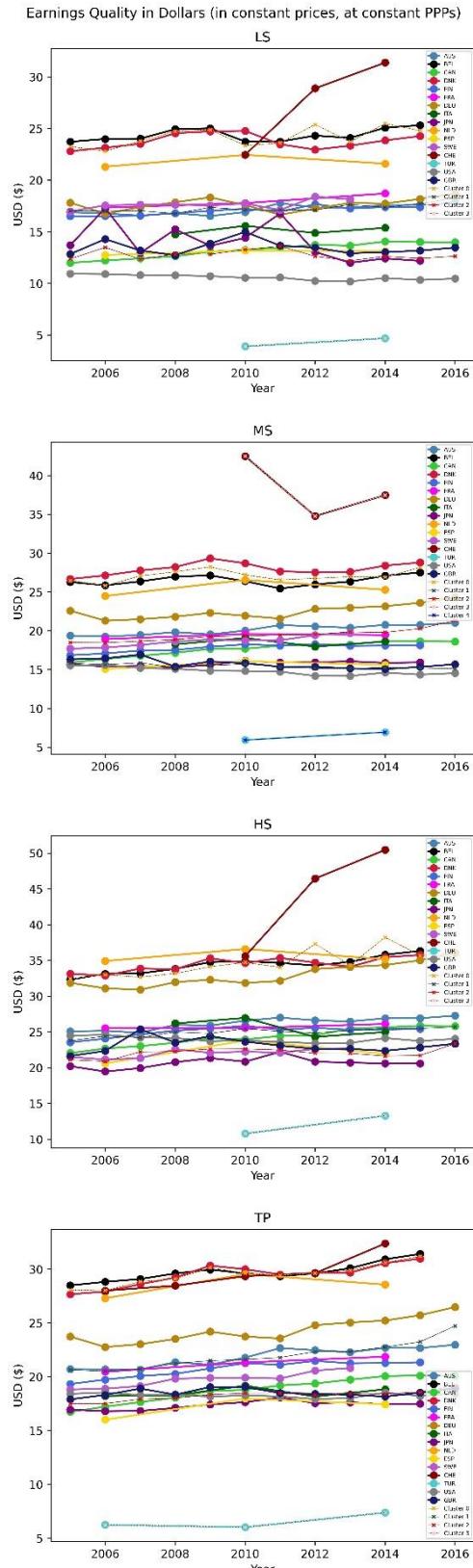
cluster	countries
Cluster 0	['BEL', 'DNK', 'NLD']
Cluster 1	['JPN', 'ESP', 'USA', 'GBR']
Cluster 2	['AUS', 'CAN', 'FIN', 'FRA', 'DEU', 'ITA', 'SWE']
Cluster 3	['CHE']
Cluster 4	['TUR']

Clusters for High Skilled:

cluster	countries
Cluster 0	['BEL', 'DNK', 'DEU', 'NLD', 'CHE']
Cluster 1	['AUS', 'CAN', 'FIN', 'FRA', 'ITA', 'USA']
Cluster 2	['JPN', 'ESP', 'SWE', 'GBR']
Cluster 3	['TUR']

Clusters for Total:

cluster	countries
Cluster 0	['BEL', 'DNK', 'NLD', 'CHE']
Cluster 1	['AUS', 'FIN', 'FRA', 'DEU', 'SWE']
Cluster 2	['CAN', 'ITA', 'JPN', 'ESP', 'USA', 'GBR']
Cluster 3	['TUR']



In general, the outstanding data are:

- When we look at the income levels of people at every skill level, we see that Turkey is by far the lowest. In other words, the income of employees in Turkey is very low, regardless of their skill level. This causes foreign countries to see Turkey as a cheap labor force.
- Belgium, Denmark, and the Netherlands are mostly in the same cluster and at the top of every skill level.
- People in Switzerland have a good income at all skill levels. Especially medium skilled people earn much more than those in other countries.

Education level is not the only factor affects earning, factors such as the socio-economic status of the country should also be considered (Tamborini,2015). However, when we look at the distribution of earnings according to education level, we can clearly see that higher education brings more income. But we see that there is not much difference in the earnings of individuals with medium and low education in Turkey. In addition, the earnings of employees in Turkey are extremely low compared to other countries. Considering that this graph only shows data up until 2016, and also taking into account the significant depreciation of Turkey's currency, the earnings obtained in Turkey based on the level of education have decreased to much more lower levels compared to other countries. Therefore, there is a significant availability of cheap labor in Turkey, making it quite reasonable for companies to open factories in Turkey.

e. Foreign Students Rate

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Education and Training" section with "Share of international students and all students by field" title.

These are all titles in this dataset:

"COUNTRY", "Country", "INDICATOR", "Indicator", "SEX", "Gender", "EDUCATION _LEV", "Education level", "MOBILITY", "Mobility", "EDUCATION_FIELD", "Field", "YEAR", "Year", "Value", "Flag Codes", "Flags"

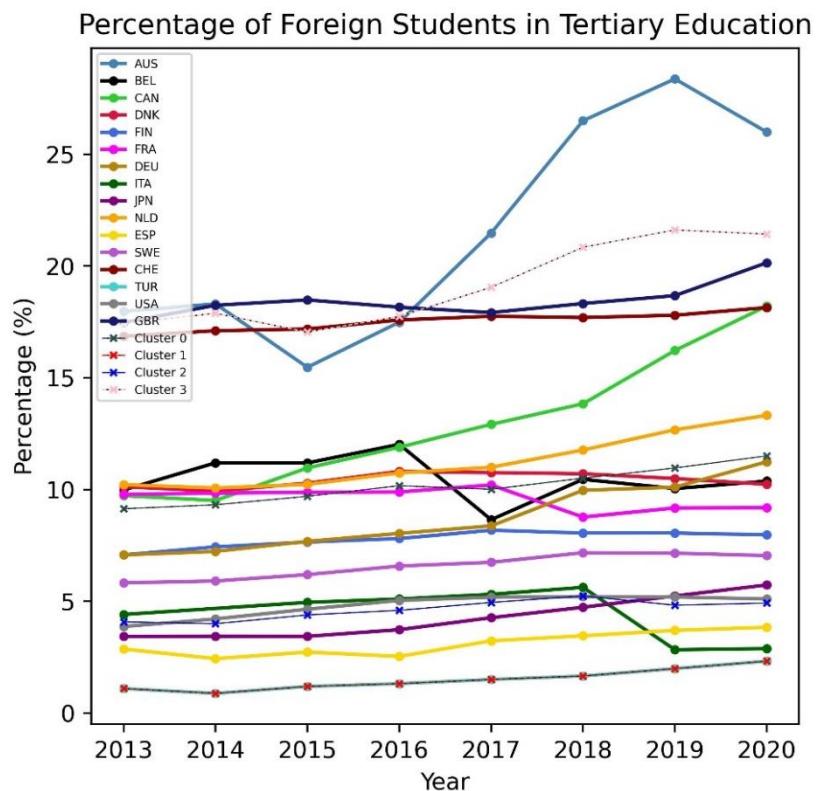
The rate of foreign students studying in a country indicates that people in other countries trust that country's education system.

This does not always mean that the education system in that country is very good, because the reason for this high rate may be due to the high number of tourists coming from abroad or living in that country temporarily.

However, it can still be informative about the level of the education system of that country.

Foreign Students in Tertiary Education

cluster	countries
Cluster 0	['BEL', 'CAN', 'DNK', 'FIN', 'FRA', 'DEU', 'NLD']
Cluster 1	['TUR']
Cluster 2	['ITA', 'JPN', 'ESP', 'SWE', 'USA']
Cluster 3	['AUS', 'CHE', 'GBR']

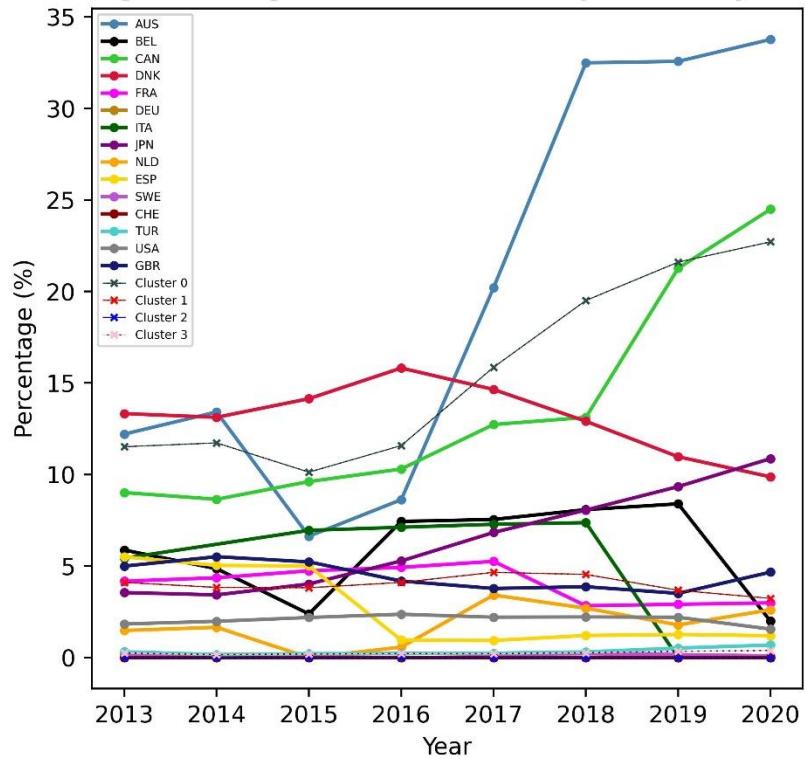


- By looking at the cluster table and chart above, we can say that Türkiye is different from the whole table. In the chart, we see that Türkiye is at a low level and has a slowly increasing chart.
- Australia, Switzerland, and Great Britain (Cluster 3) on top of the chart.

Short-Cycle Tertiary Education:

cluster	countries
Cluster 0	['AUS', 'CAN', 'DNK']
Cluster 1	['BEL', 'FRA', 'ITA', 'JPN', 'NLD', 'ESP', 'SWE', 'TUR', 'USA', 'GBR']
Cluster 2	['DEU', 'CHE']

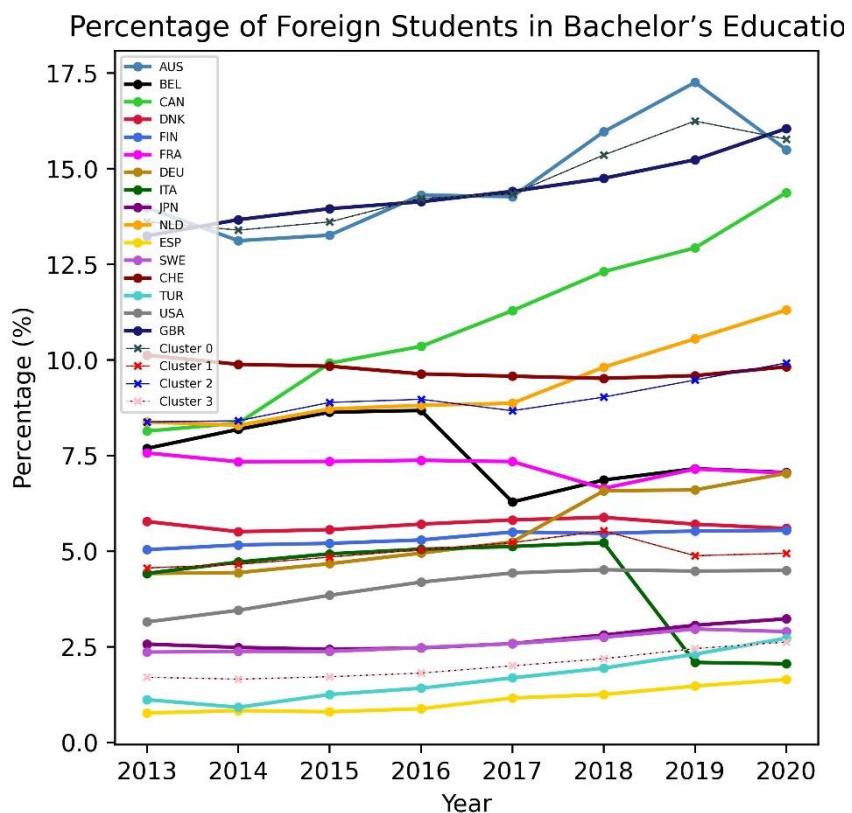
Percentage of Foreign Students in Short-cycle Tertiary Education (%)



- There seems to be no Short-Cycle Tertiary Education in Germany and Switzerland (Cluster 2).
- Australia is increasing the percent of Short-Cycle Tertiary Education.

Bachelor's Education:

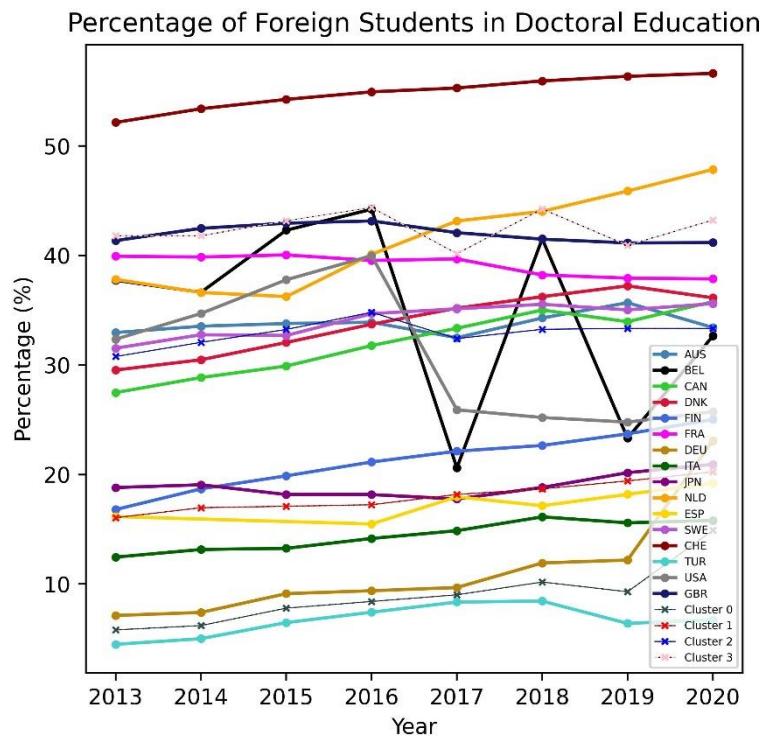
cluster	countries
Cluster 0	['AUS', 'GBR']
Cluster 1	['DNK', 'FIN', 'DEU', 'ITA', 'USA']
Cluster 2	['BEL', 'CAN', 'FRA', 'NLD', 'CHE']
Cluster 3	['JPN', 'ESP', 'SWE', 'TUR']



- Great Britain and Australia (Cluster 0) at the top of the chart.
- Japan, Spain, Sweden, and Türkiye (Cluster 3) at the bottom of the chart as the lowest cluster.

Doctoral Education:

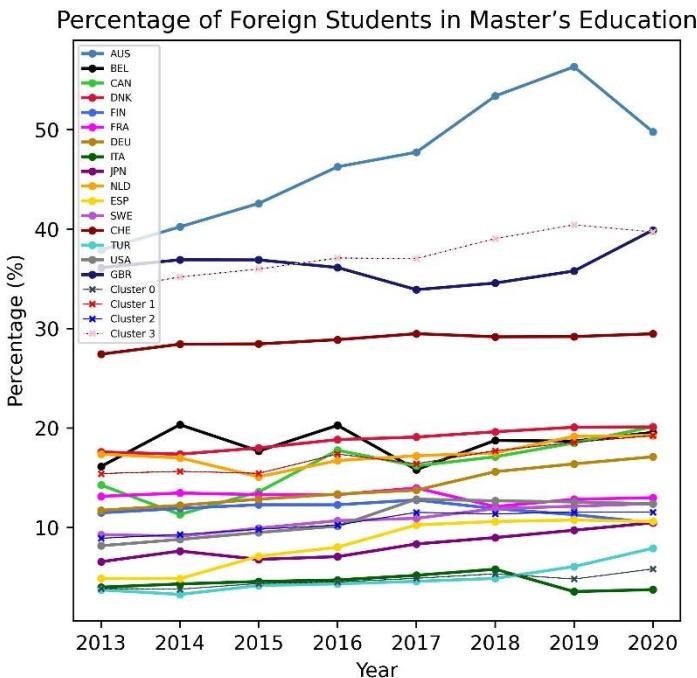
cluster	countries
Cluster 0	['DEU', 'TUR']
Cluster 1	['FIN', 'ITA', 'JPN', 'ESP']
Cluster 2	['AUS', 'CAN', 'DNK', 'SWE', 'USA']
Cluster 3	['BEL', 'FRA', 'NLD', 'CHE', 'GBR']



- Germany and Türkiye (Cluster 3) at the bottom of the chart as the lowest Cluster.

Master's Education:

cluster	countries
Cluster 0	['ITA', 'TUR']
Cluster 1	['BEL', 'CAN', 'DNK', 'DEU', 'NLD']
Cluster 2	['FIN', 'FRA', 'JPN', 'ESP', 'SWE', 'USA']
Cluster 3	['AUS', 'CHE', 'GBR']



- We see Australia, Switzerland, and Great Britain (Cluster 3) at the top of the chart.
- We see Türkiye and Italy (Cluster 0) at the bottom of the chart.

In general, the outstanding data about foreign students' rate:

- Türkiye were mostly at the bottom of the chart, which means there is no attraction to Türkiye for education.
- Australia and especially Great Britain were mostly at the top of the charts. This was expected for Great Britain, because universities such as Oxford and Cambridge are world-renowned and most successful universities.

Students migrate to another country with the aim of receiving a better education. Although this migration is mostly for educational purposes only, some choose to stay in that country. Even in countries with a high number of foreign students like the United States, the stay rate of students who go for education purposes does not exceed 50% (Spilimbergo, 2009). Therefore, it would not be wrong to assume that the majority of students go only for educational purposes. While a high rate of foreign students indicates a good education system, it does not necessarily mean that the country provides the best education in the world.

7. Safety (Security)

Security is a fundamental factor for the well-being of the people. The high crime rate in the country causes the people to feel insecure. It can also cause chaos in the country and endanger the future of the country.

The crime rate in the country and the frequency of traffic accidents show how many people feel safe people in that country.

a. Crime

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Regions and Cities" section with "Safety in regions" title.

These are all titles in this dataset:

"TL", "Territory level and

"Typology", "REG_ID", "Region", "VAR", "Indicator", "POS", "Position", "TIME", "Year", "Unit Code", "Unit", "PowerCode Code", "PowerCode", "Reference Period Code", "Reference Period", "Value", "Flag Codes", "Flags"

The constitution is essential to ensure the freedom of the people, order, and security of society. A crime is any activity prohibited by the constitution. High crime rate prevents people in that country from feeling safe, and people cannot live in safety.

We will examine some data to give an idea about crime rates:

1. Intentional Homicide

Intentional Homicide is the intentional death of someone without accident. This is a very serious crime and should have very heavy penalties in the constitution.

Since Intentional Homicide is one of the most serious crimes, it can be informative about how big crimes are committed in that country.

cluster	countries
Cluster 0	['BEL', 'CAN']
Cluster 1	['DNK', 'FIN', 'DEU', 'ITA', 'JPN', 'NLD', 'ESP', 'SWE', 'CHE']
Cluster 2	['TUR']
Cluster 3	['USA']
Cluster 4	['AUS', 'FRA', 'GBR']

We see Türkiye and USA are alone in their clusters. We see Belgium and Canada in same cluster (Cluster 0).

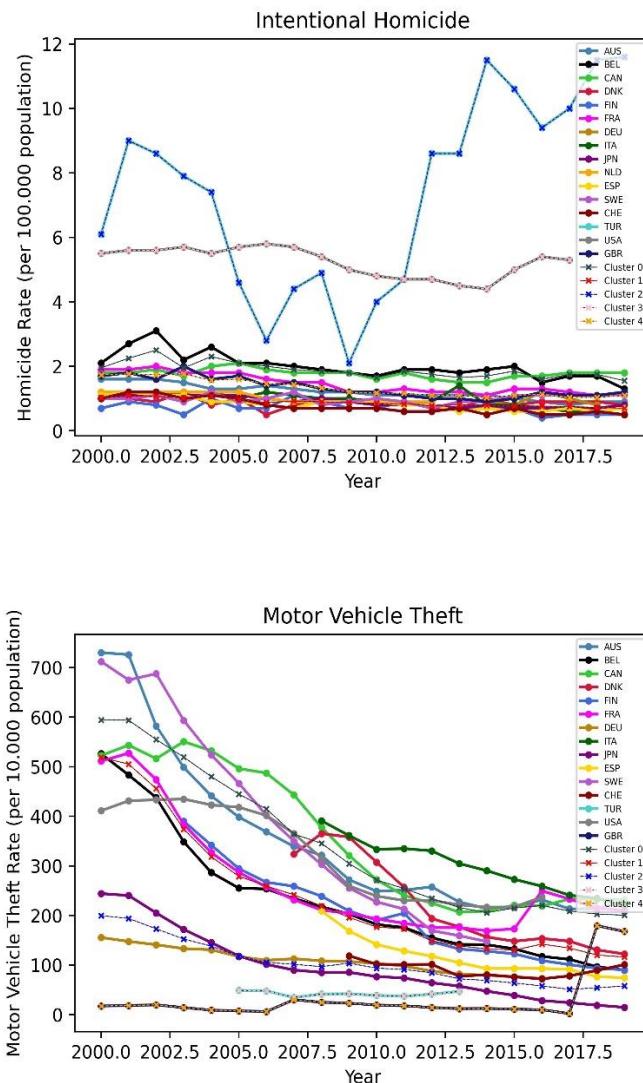
2. Motor Vehicle Theft

Motor Vehicle Theft is not a petty theft crime. This is a major crime and should have severe penalties.

Motor Vehicle Theft can be informative in terms of thefts in that country.

cluster	countries
Cluster 0	['AUS', 'CAN', 'DNK', 'ITA', 'SWE', 'USA']
Cluster 1	['BEL', 'FIN', 'FRA', 'ESP']
Cluster 2	['DEU', 'JPN', 'CHE']
Cluster 3	['TUR']
Cluster 4	['GBR']

Crime Data



In general, the outstanding data about crime are:

- The crime of “Intentional Homicide” is much higher in Türkiye and the USA than in any other country. However, the data of these two countries are not similar. While the USA stays at the top level more steadily, Turkey is more unstable and can go very high and low.
- “Motor Vehicle Theft” is a declining crime in most of all countries over the years.
- Nevertheless, although “Intentional Homicide” crime is high and unstable in Turkey, we see that this crime rate is low.
- While it has decreased in many countries, we see that this crime has increased in Great Britain after 2017.

The crime rate can be attributed to many factors such as education, economy, mental health. In societies with high income inequality, the crime rate is high because the polarization and hatred in the society is high. Nadanovsky also conducted studies showing that as income inequality in the society increases, the crime rate increases (2009). According to Lochner, the probability of committing a crime decreases as the education level of individuals increases (2020). Therefore, the government should attach great importance to the education of the society. In addition, since socialization is very important for mental health, it is very important for people to find time and money to socialize.

b. Traffic Accident

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Regions and Cities" section with "Safety" title.

These are all titles in this dataset:

"COUNTRY","Country","INDICATOR","Indicator","YEAR","Year","Value","Flag Codes","Flags"

Traffic Accidents are unintentional situations. These accidents can have many reasons such as lack of attention of the driver, bad road, car breakdown. Although these accidents occurred unintentionally, they are situations that endanger the safety of people.

Road Fatalities per 1.000.000 vehicle-km:

cluster	countries
Cluster 0	['AUS', 'CAN', 'FIN', 'DEU', 'NLD', 'CHE']
Cluster 1	['FRA', 'JPN', 'USA']
Cluster 2	['BEL', 'ESP']
Cluster 3	['TUR']
Cluster 4	['DNK', 'SWE', 'GBR']

We see Türkiye alone in its cluster.

Belgium and Spain (Cluster 2) together in same cluster.

Road Fatalities Rate by Inhabitants:

cluster	countries
Cluster 0	['AUS', 'CAN', 'FIN', 'TUR']
Cluster 1	['BEL', 'FRA', 'ITA', 'ESP']
Cluster 2	['DNK', 'DEU', 'JPN', 'NLD', 'CHE']
Cluster 3	['SWE', 'GBR']
Cluster 4	['USA']

In this dataset, we see the USA alone in its cluster.

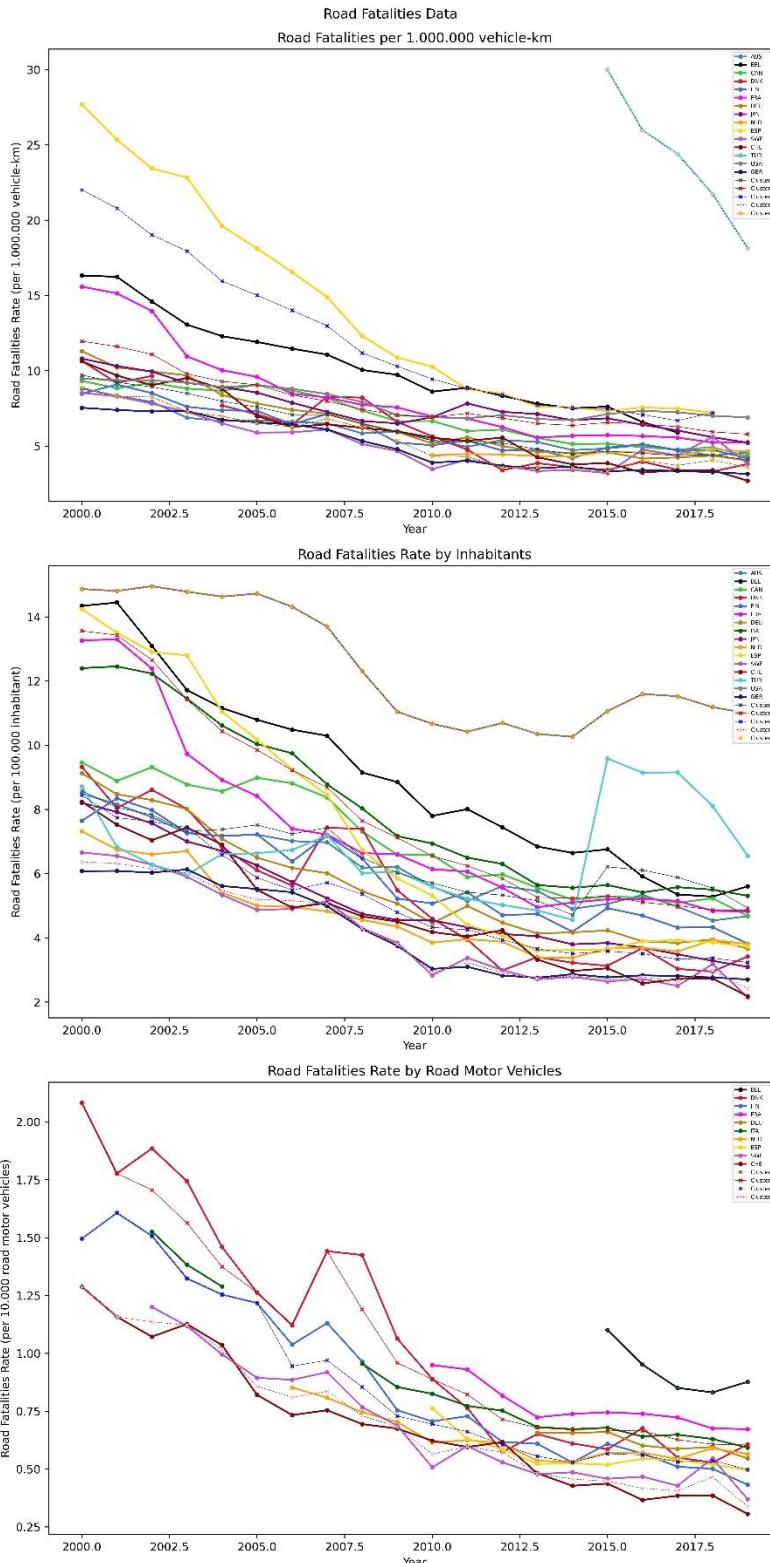
Sweden and Great Britain (Cluster 3) together in same cluster.

Road Fatalities Rate by Road Motor Vehicles:

cluster	countries
Cluster 0	['BEL']
Cluster 1	['DNK', 'FRA', 'DEU', 'ITA']
Cluster 2	['FIN', 'NLD', 'ESP']
Cluster 3	['SWE', 'CHE']

Belgium is an outlier in this dataset.

Sweden and Switzerland (Cluster 3) together in same cluster.



In general, the outstanding data about traffic accident rates are:

- “Road Fatalities per 1.000.000 vehicle-km” are very common in Türkiye. Although it is in a downward trend like other countries, the gap is very huge.
- “Road Fatalities rate by inhabitants” are very high in USA.
- “Road Fatalities” per 1.000.000 vehicle-km was very high and there was a downward trend for Türkiye in years between 2015-2018, we see that “rate by inhabitants” is also decreasing in this chart. Surprisingly, Turkey is not at the top here. This may be due to the fact that people who do not travel much are also included.
- Belgium is different from other countries and is in a separate cluster with its high rates.

Traffic accident rate can be considered in conjunction with various factors such as the overall vehicle usage rate in the country and early-age driving. Wegman believes that although not entirely certain, the economic decline has contributed to reducing the mortality rate of accidents among young drivers by reducing alcohol consumption (2017). It can be expected that the accident rate would decrease due to the presence of an economic decline, which would decrease factors such as vehicle usage and alcohol consumption. However, economic hardships can increase factors such as stress and the number of faulty vehicles, which may prevent a decrease in the accident rate. Since more detailed research and data are needed to be certain, it is not possible to reach a definitive conclusion with this data.

8. Social Connections

As Aristotle said, "Man is a social animal by nature". In order for people to have a healthy mind, they must have a social life. Citizens should be able to spare time for social life and socialize by spending time with each other.

a. Social Connections

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Education and Training" section with "Social connections" title.

These are all titles in this dataset:

"COUNTRY", "Country", "ISC11A", "ISCED-A
2011", "SEX", "Gender", "AGE", "Age", "PIAAC_CATEGORY", "Category", "INDICATOR", "Indicator", "MEASURE", "Measure", "YEAR", "Reference year", "Value", "Flag Codes", "Flags"

Social Life is indispensable for people. It is also essential for mental relaxation. Let's take a look at what people in OECD countries are saying about their social lives:

Adults who participated in any cultural or sporting activities in the last 12 months:

cluster	countries
Cluster 0	['DNK', 'FIN', 'FRA', 'NLD', 'SWE', 'GBR']
Cluster 1	['ITA']
Cluster 2	['BEL', 'DEU', 'ESP']

We see Italy as an outlier in these clusters.

Adults who get together with friends living outside their household at least once a week:

cluster	countries
Cluster 0	['BEL', 'FIN', 'ITA', 'ESP', 'SWE']
Cluster 1	['DNK']
Cluster 2	['FRA', 'DEU']
Cluster 3	['NLD', 'GBR']

We see Denmark as an outlier in these clusters.
 France and Germany have similar values, and also Netherlands and Great Britain have similar values.

Adults who have someone to ask for help:

cluster	countries
Cluster 0	['BEL', 'DNK', 'FIN', 'FRA', 'DEU', 'NLD', 'ESP', 'SWE', 'GBR']
Cluster 1	['ITA']

Italy is the only different country in this field.

Adults who actively participate in social media on a daily basis:

cluster	countries
Cluster 0	['BEL', 'ESP']
Cluster 1	['DNK', 'FIN', 'NLD', 'SWE', 'GBR']
Cluster 2	['FRA']
Cluster 3	['DEU', 'ITA']

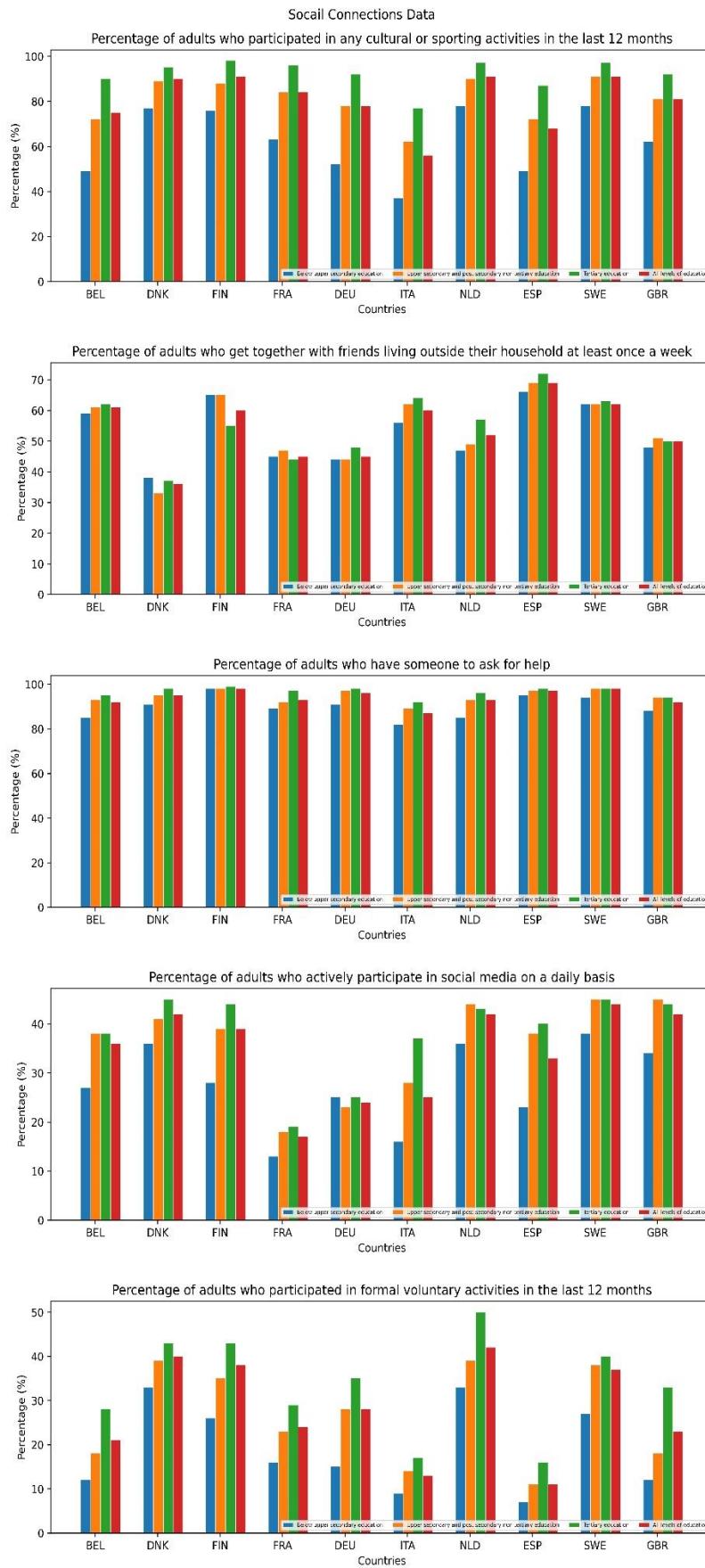
We see France as an outlier in these clusters.
 Belgium and Spain have similar values, and also Germany and Italy have similar values.

Adults who participated in formal voluntary activities in the last 12 months:

cluster	countries
Cluster 0	['DNK', 'FIN', 'NLD', 'SWE']
Cluster 1	['ITA', 'ESP']
Cluster 2	['BEL', 'FRA', 'DEU', 'GBR']

Countries are divided into 3 clusters.
 We see Italy and Spain has similar values.

Let's see our chart:



In general, the outstanding data about Social Connections Data are:

- We see that people do not participate in cultural or sports activities very often in Italy.
- We see that more than half of the people participate in cultural or sports activities very often in Belgium, Germany, and Spain.
- People don't go out much to socialize in Denmark.
- France and Germany have similar values which is around 40% and they don't like so much hanging out.
- People are not afraid to ask someone for help in all countries. Italy is a little bit below than other countries, so it seems like an outlier in clusters.
- People in France do not like spending time on social media.
- Germany and Italy have similar values and they don't really like spending time on social media.
- Belgium and Spain have similar values and some of the people like spending time on social media.
- People in Italy and Spain do not participate in formal voluntary activities much.

Since social and emotional skills affect mental and physical life, it is one of the most important factors affecting quality of life (Chernyshenko, 2018). Since quality of life is a variable that affects very important factors such as subjective well-being and health, it is an important indicator to measure the well-being of the country. Since it would not be very rational to measure emotional skills, we can see which societies are more social according to their social life preferences, thanks to this data. In general, we can say that the social lives of countries are close to each other, except for a few data.

b. Social Spending

This dataset is retrieved from “OECD Data” as ".csv" file. This dataset can be reached from <https://data.oecd.org/socialexp/social-spending.htm> with “Social spending” title. These are all titles in this dataset:

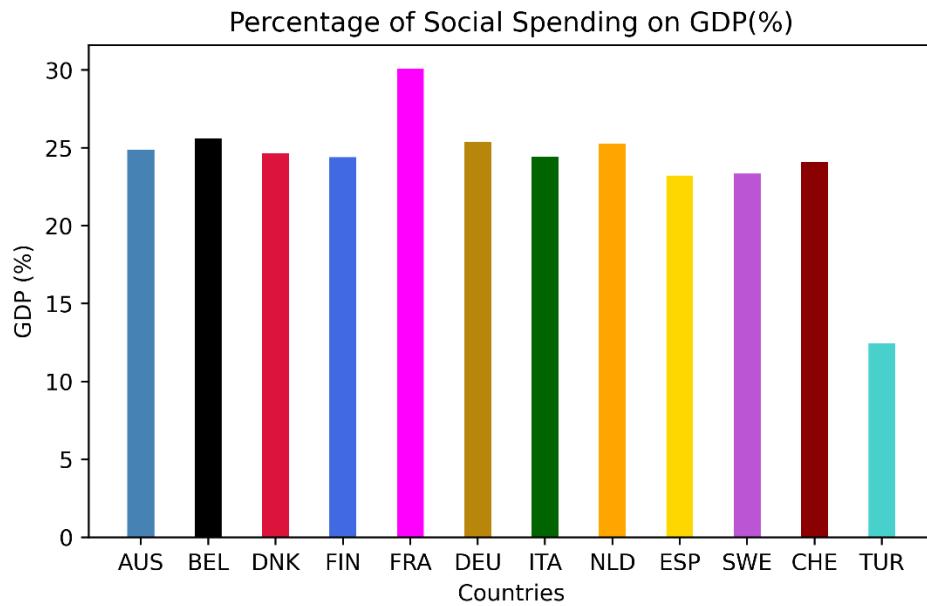
“LOCATION”, “INDICATOR”, “SUBJECT”, “MEASURE”, “FREQUENCY”, “TIME”, “Value”, “Flag Codes”

The state should spend some of the money for social purposes. This expenditure should not be considered unnecessary. Because the money spent for social expenditures is very important for the welfare of the country. Examples of these expenditures are pension payments, health expenditures, and cash aid to poor families. This data is important in terms of understanding the importance that the country attaches to social equality.

It helps us understand how much of the income is spend for public.

cluster	countries
Cluster 0	['FRA']
Cluster 1	['BEL', 'DEU', 'NLD']
Cluster 2	['AUS', 'DNK', 'FIN', 'ITA', 'CHE']
Cluster 3	['ESP', 'SWE']
Cluster 4	['TUR']

France and Türkiye can be assumed to be the outliers in this clusters.



- The money spent for social expenses in Turkey is half as much as in other countries. The following can be said about this:
 - The government may have reduced social spending as it was going through an economic depression. Although this is not a correct policy, it may be the reason why this value is not low.
 - In Turkey, where income inequality is already high, the low level of social expenditures will further increase this inequality.
- The country that spends the most money is France. Even if the income of people in France is not very high, the state helps them. This is a very correct policy for the understanding of a social country.

The importance that countries place on social understanding can be understood by their percentage of GDP. If the public can't spend money on socializing or there are situations such as income inequality, governments need to help the poor more to reduce this inequality and make sure the balance. According to Adema's research, most of the money in this sector is spent on health and retirement (2011). In the same study, social spending increased by about 20 percent from 1980 to 2008. If we consider inflation, countries have not increased their social expenditures as much as necessary. For this reason, it is essential to give the necessary importance to social expenditures for the reduction of income inequality and the welfare of the country.

9. Work and Job Quality

People need to have an income in order to continue their lives, and they need to have a job to make a living. Having a job not only makes people money, but it also allows people to socialize and improve their skills.

Employment Rates and Unemployment Rates are very important to have a richer country. Also, long-term unemployment shows how long people wait to find a job, and youth unemployment rate shows how easy or hard it is for young people to find work. In addition, the quality of the work allows us to measure whether people trust their work and the job strain allows us to measure how much and how hard people work.

a. Employment

The Employment Rate indicates what percentage of the labour resources (people that can work) in the country is used. This ratio allows us to comment on the economic system and productivity in country. In this way, it can be commented on the level of welfare in the country.

i. Employment Rate by Gender

This dataset is retrieved from “OECD Data” as ".csv" file. This dataset can be reached from <https://data.oecd.org/emp/employment-rate.htm> with “Employment Rate” title.

These are all titles in this dataset:

"LOCATION", "INDICATOR", "SUBJECT", "MEASURE", "FREQUENCY", "TIME", "Value", "Flag Codes"

This rate shows the distribution of people working in the country by gender. In this way, we can see in which country women and men employees are more balanced.

Men:

cluster	countries
Cluster 0	['DNK', 'ITA', 'USA']
Cluster 1	['AUS', 'BEL', 'CAN', 'FIN', 'FRA', 'NLD', 'CHE']
Cluster 2	['DEU', 'JPN', 'SWE', 'GBR']
Cluster 3	['ESP']
Cluster 4	['TUR']

The rate of working men is more different in Spain and Turkey than in other countries.

Women:

cluster	countries
Cluster 0	['AUS', 'FIN', 'FRA', 'ITA', 'ESP', 'SWE', 'CHE', 'GBR']
Cluster 1	['JPN']
Cluster 2	['BEL', 'DEU', 'NLD']
Cluster 3	['TUR']
Cluster 4	['CAN', 'DNK', 'USA']

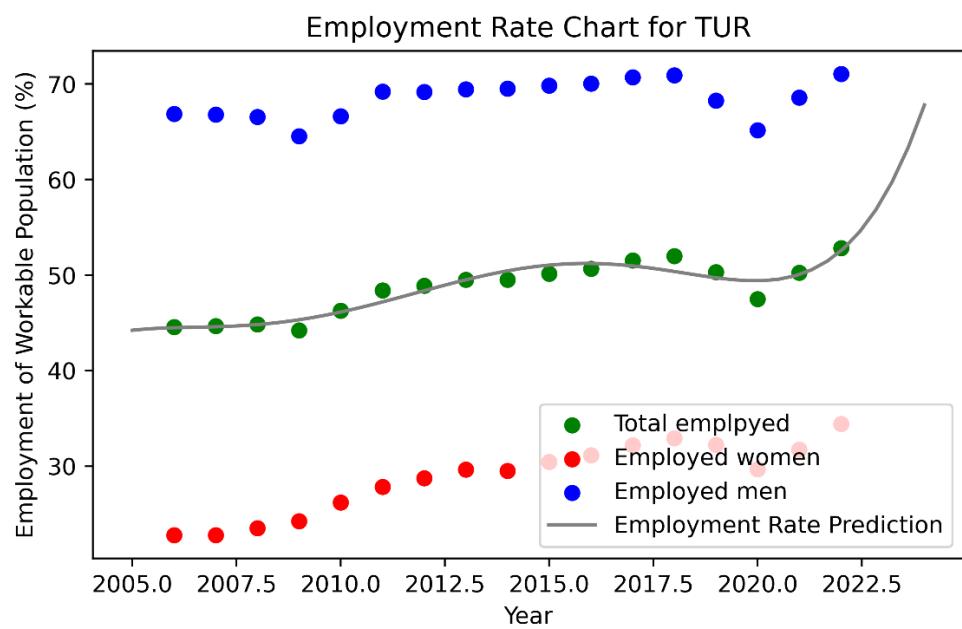
The rate of working women is more different in Japan and Turkey than in other countries.

Total:

cluster	countries
Cluster 0	['AUS', 'BEL', 'CAN', 'FRA', 'SWE', 'CHE', 'GBR']
Cluster 1	['DEU', 'JPN', 'NLD']
Cluster 2	['DNK', 'FIN', 'ITA', 'USA']
Cluster 3	['ESP']
Cluster 4	['TUR']

As with the male employment rate, we see that the total employment rate in Spain and Turkey is different than in other countries.

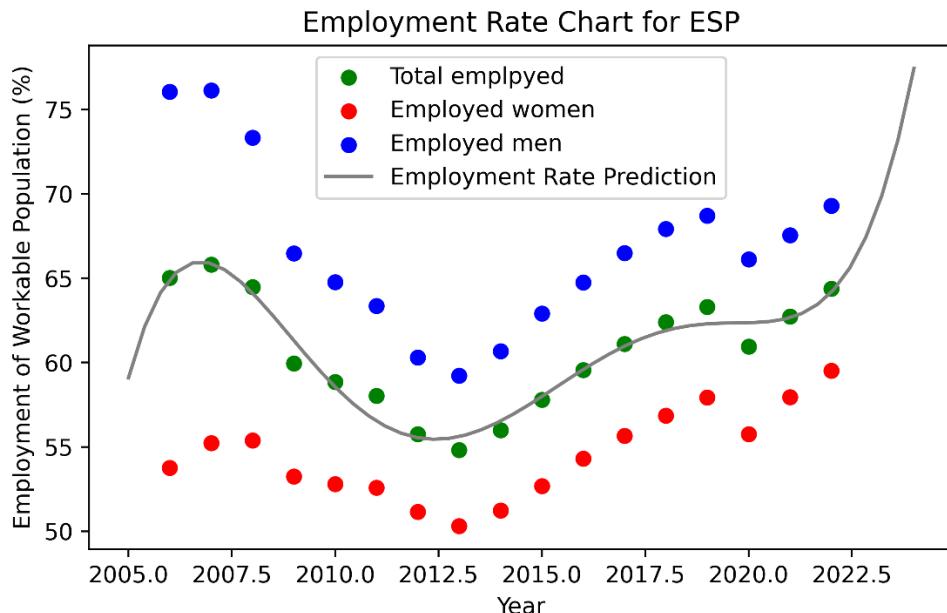
Türkiye



R2 score:0.9

- The gap between the employment rates of female and male is much higher than in other countries. Although this difference has decreased somewhat over the years, it is still huge.
- Although the employment rate has increased slightly, this rate is extremely low compared to other countries.
- We see that the decrease in the employment rate with COVID-19 is also in Turkey, as in every country.

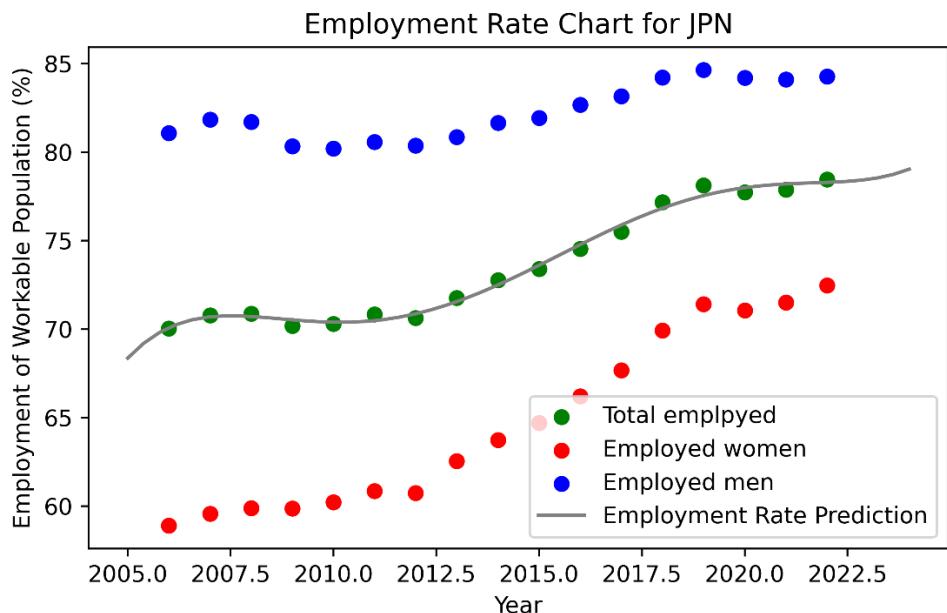
Spain



R2 score:0.95

- Increases and decreases in the male and total employment rate of Spain are not similar to any other country. Therefore, the algorithm clustered this country separately.
- The gap in the employment rate of men and women in the country has gradually narrowed.

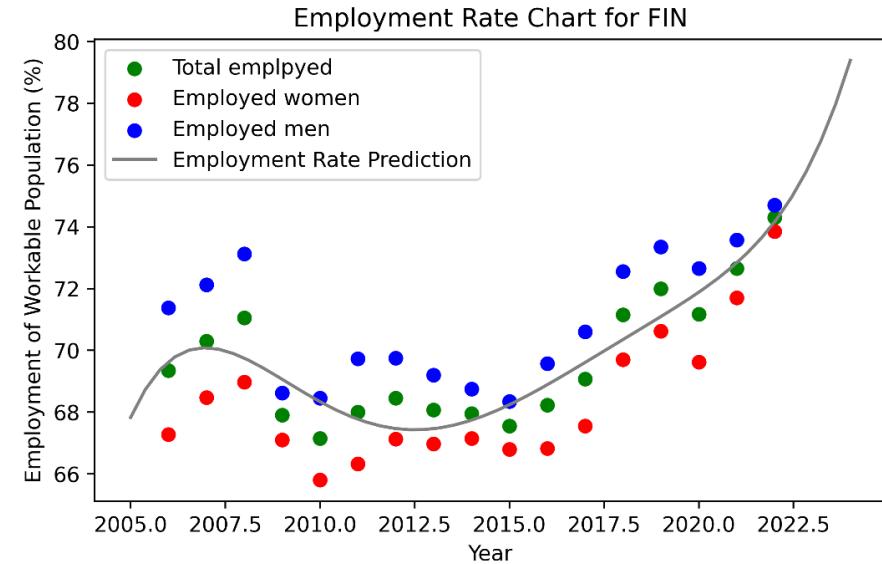
Japan



R2 score:0.99

- Female, male, and total employment rate is higher than in many countries.
- Japan ranks in the middle in employment rate.

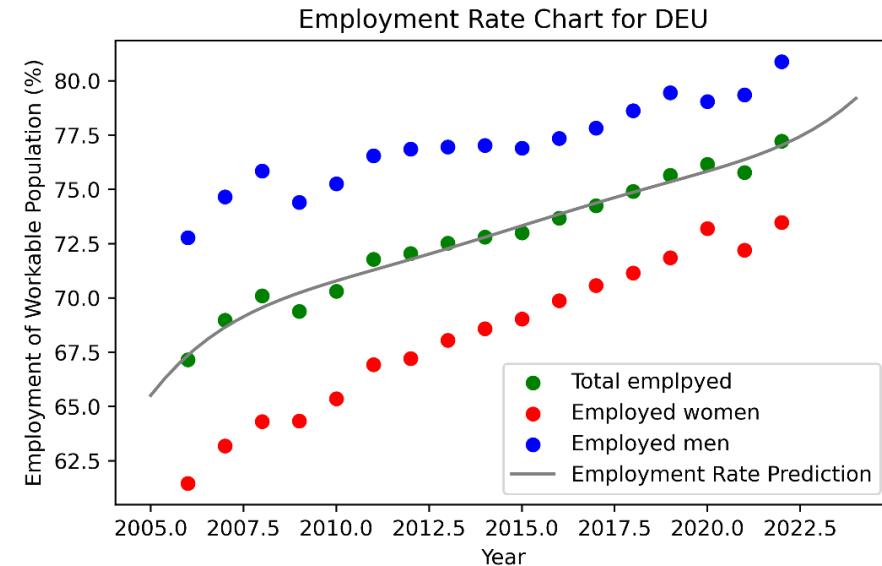
Finland



R2 score: 0.87

- The female and male employment rate in Finland has been close from the beginning. Especially in recent years, these rates are almost equal to each other. It would not be wrong to say that there is gender equality in the country.
- The country's total employment rate is also quite high.

Germany



R2 score: 0.98

- The gap between female and male employment rates in Germany has steadily decreased.
- The country has significantly and steadily increased its employment rate.
- Although the employment rate is constantly increasing, we see a slight decrease in male and total employment rates in 2020.
- Surprisingly, this year, however, the employment rate of women has increased. This may be due to the right policies implemented by the government.

ii. Employment by Education Level

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Education and Training" section with "Employment rate, by level of educational attainment and age group" title.

These are all titles in this dataset:

"LOCATION", "INDICATOR", "SUBJECT", "MEASURE", "FREQUENCY", "TIME", "Value", "Flag Codes"

In this data set, we will examine the relationship between the employment rate and the education level of people.

Below upper Secondary:

cluster	countries
Cluster 0	['AUS', 'BEL', 'FIN', 'FRA']
Cluster 1	['CAN', 'DNK', 'ITA', 'NLD', 'SWE', 'CHE', 'USA', 'GBR']
Cluster 2	['ESP']
Cluster 3	['DEU', 'TUR']

Spain is in a separate cluster at this level of education.

Germany and Turkey, which have very different employment rates, are in the same cluster at this education level.

Upper secondary non-Tertiary:

cluster	countries
Cluster 0	['AUS', 'BEL', 'CAN', 'FIN', 'FRA', 'ITA', 'NLD']
Cluster 1	['DEU']
Cluster 2	['ESP']
Cluster 3	['DNK', 'SWE', 'CHE', 'TUR', 'USA', 'GBR']

Spain is also in a separate cluster at this level of education.

Germany, which has a high employment rate, has its own cluster here.

Tertiary:

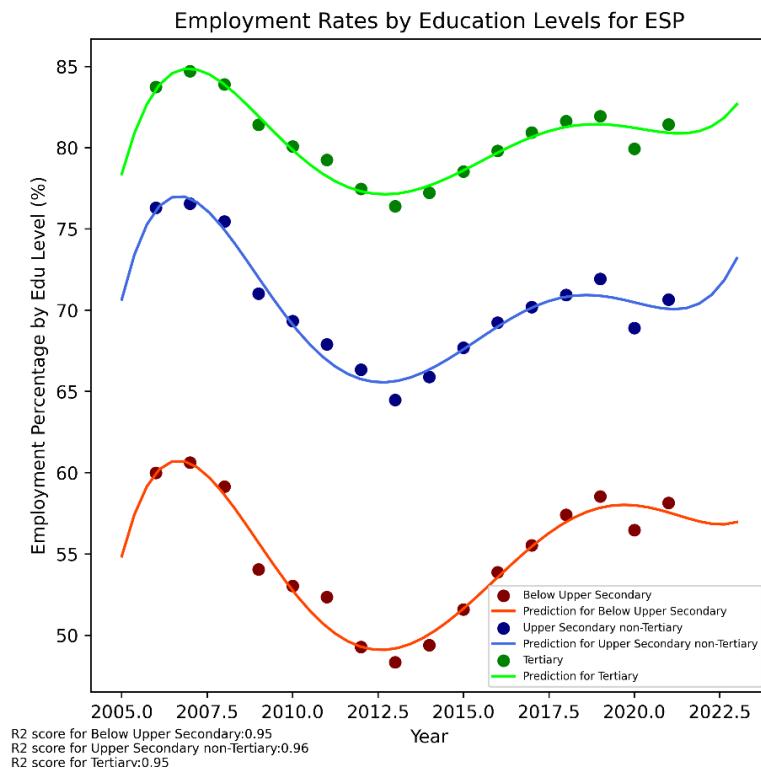
cluster	countries
Cluster 0	['AUS', 'CAN']
Cluster 1	['DNK', 'FIN', 'ITA', 'CHE', 'USA', 'GBR']
Cluster 2	['BEL', 'FRA', 'DEU', 'NLD', 'SWE']
Cluster 3	['ESP']
Cluster 4	['TUR']

Spain is also in a separate cluster at this level of education.

Türkiye, which has a low employment rate compared to other countries, has its own cluster here.

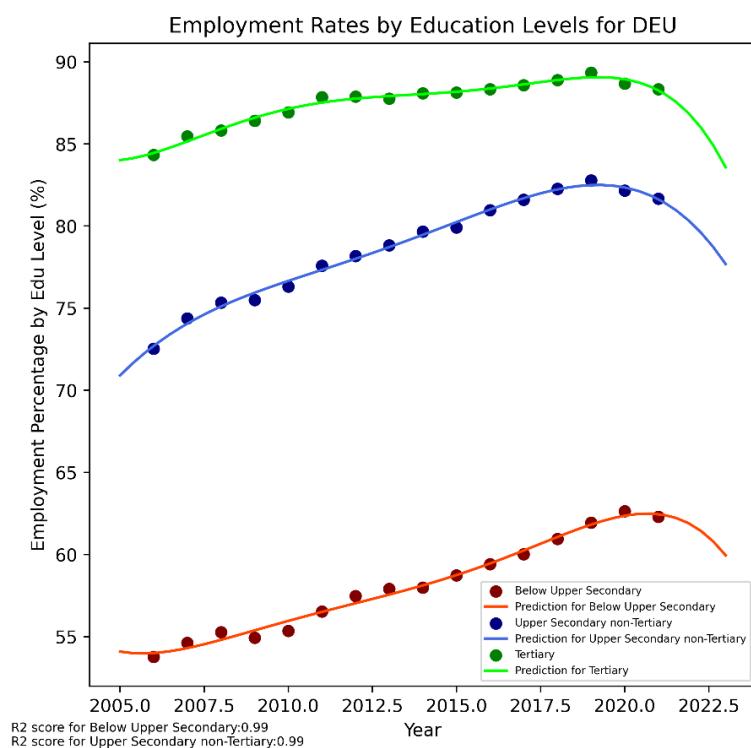
Australia and Canada are in the same cluster together.

Spain



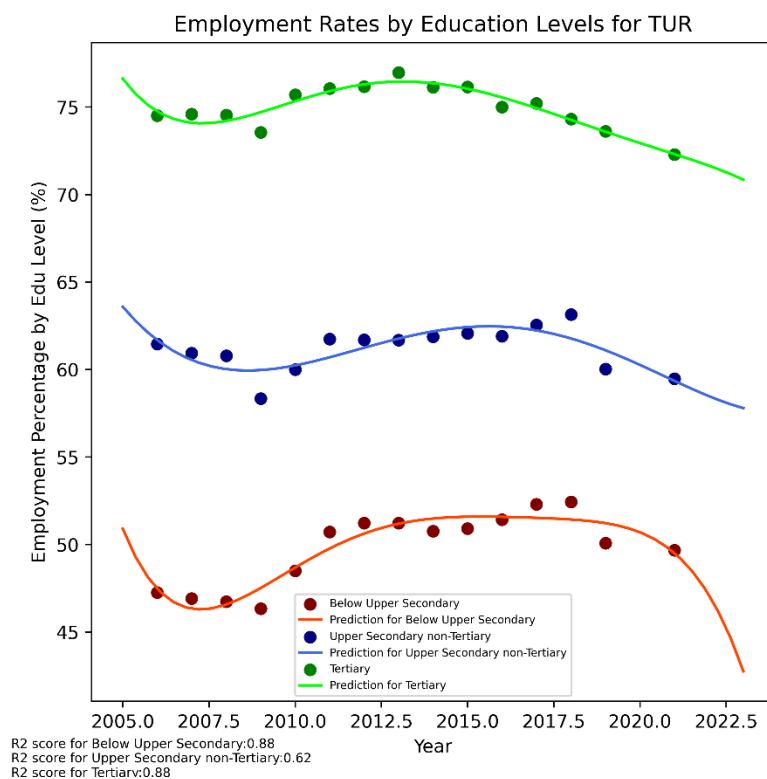
- Spain's chart is not like any other country. Especially in 2012, the decline in every level of education is not so evident in any country.

Germany



- Germany maintains its steady growth at all educational levels.
- The distribution of educational levels is very regular.

Türkiye



- We see a decrease in tertiary and upper secondary non-Tertiary levels. There is no other country showing such a decline and this is quite dangerous for Turkey's future.
- Although it seems more stable for the below upper secondary, we can say that it is also in a downward trend based on the prediction graph which is quite accurate.

The employment rate is an indicator of a working country. A society that does not produce cannot develop. Policies implemented by governments are also important in increasing the employment rate. He showed how correct he could be by drawing attention to the policies implemented to increase the employment rate after the Great Depression (Immervoll, 2012). The right policies implemented by governments will benefit both employers and employees, and the employment rate will increase.

b. Unemployment

The unemployment rate helps us to see many problems such as the economic crisis and production problems in that country. It is not enough to just look at the unemployment rate for this.

The long-term unemployment rate, the rate of the young unemployed, and the unemployment rate of people according to their education level also provide information about the current and future situation of the country.

i. Unemployment Rate

This dataset is retrieved from “OECD Data” as ".csv" file. This dataset can be reached from <https://data.oecd.org/unemp/unemployment-rate.htm> in with “Unemployment Rate” title.

These are all titles in this dataset:

"LOCATION", "INDICATOR", "SUBJECT", "MEASURE", "FREQUENCY", "TIME", "Value", "Flag Codes"

The unemployment rate can be explained as unused labour force.

Generally, people cannot work because they cannot find a job, and this is a sign that there are problems in that country.

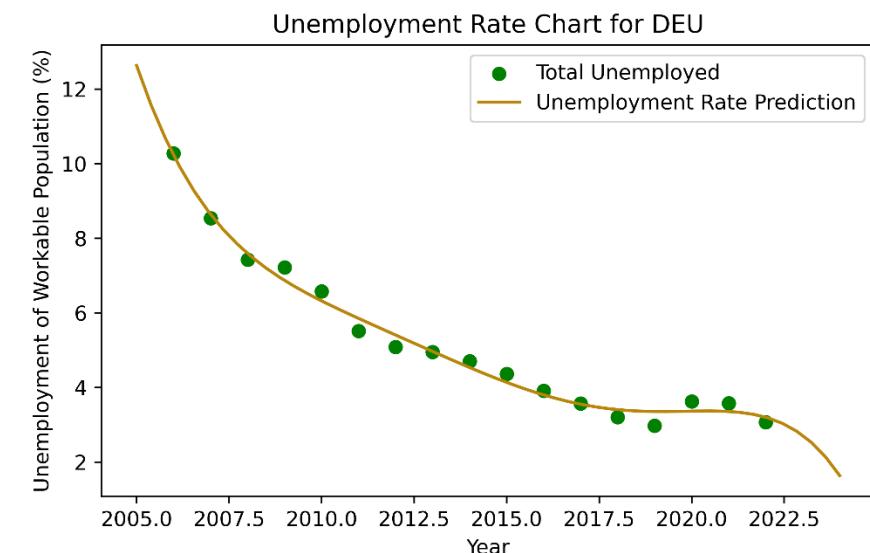
The unemployment rate cannot give us detailed information, but it is sufficient to give information about the general situation of the country.

Since the data provided by OECD does not include the distribution of unemployment rate by gender, we will not be able to comment on the unemployment rate by gender in this data set.

cluster	countries
Cluster 0	['DNK', 'ESP']
Cluster 1	['BEL', 'FIN', 'FRA', 'NLD']
Cluster 2	['DEU']
Cluster 3	['AUS', 'CAN', 'ITA', 'SWE', 'CHE', 'TUR']
Cluster 4	['JPN', 'USA', 'GBR']

We remember, Germany has high employment rates. We expect unemployment rates are low for Germany.
Denmark and Spain are in same cluster.

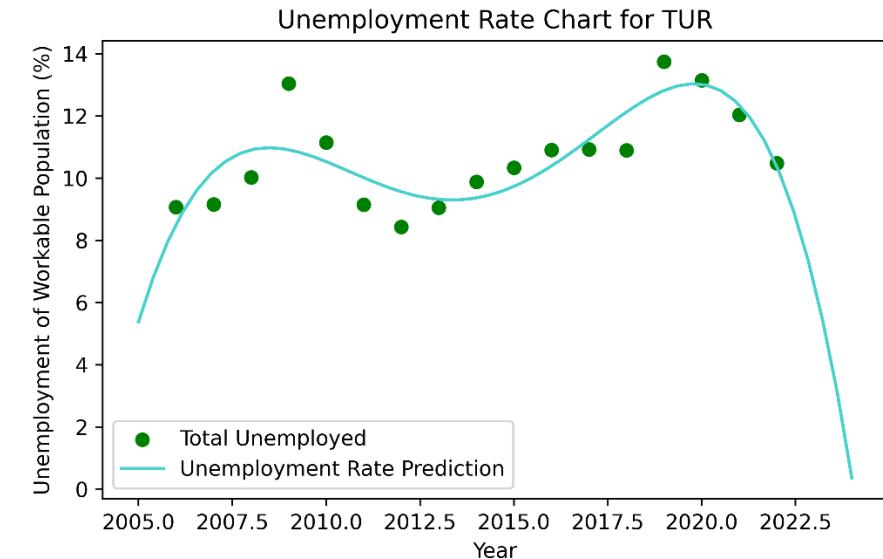
Germany



R2 score: 0.99

- Germany had a constantly increasing employment rate. In this chart, we see that there is a constantly falling unemployment rate. We can say that Germany has implemented the right policies in this area.

Türkiye



R2 score: 0.68

- Türkiye's employment rate was a slowly increasing graph. Here we see a more unstable but increasing unemployment rate.

ii. Unemployment by Education Level

This dataset is retrieved from “OECD Data” as ".csv" file. This dataset can be reached from <https://data.oecd.org/unemp/unemployment-rates-by-education-level.htm#indicator-chart> with “Unemployment rates by education level” title. These are all titles in this dataset:

"LOCATION", "INDICATOR", "SUBJECT", "MEASURE", "FREQUENCY", "TIME", "Value", "Flag Codes"

Unemployment by Education Level shows whether people remain unemployed when they receive education. This shows the quality of education received in that country and whether individuals are trained in the areas the sector needs.

Below Upper Secondary:

cluster	countries
Cluster 0	['DEU']
Cluster 1	['DNK', 'NLD', 'ESP']
Cluster 2	['AUS', 'BEL', 'CAN', 'FIN', 'FRA', 'ITA', 'SWE', 'CHE', 'TUR']
Cluster 3	['USA', 'GBR']

The rate of unemployed low educated people rate is different in Germany than in other countries.

Upper Secondary non-Tertiary:

cluster	countries
Cluster 0	['DNK', 'ESP', 'USA', 'GBR']
Cluster 1	['AUS', 'BEL', 'CAN', 'FIN', 'FRA', 'ITA', 'SWE', 'CHE', 'TUR']
Cluster 2	['DEU']
Cluster 3	['NLD']

Also, the rate of unemployed medium educated people rate is different in Germany.

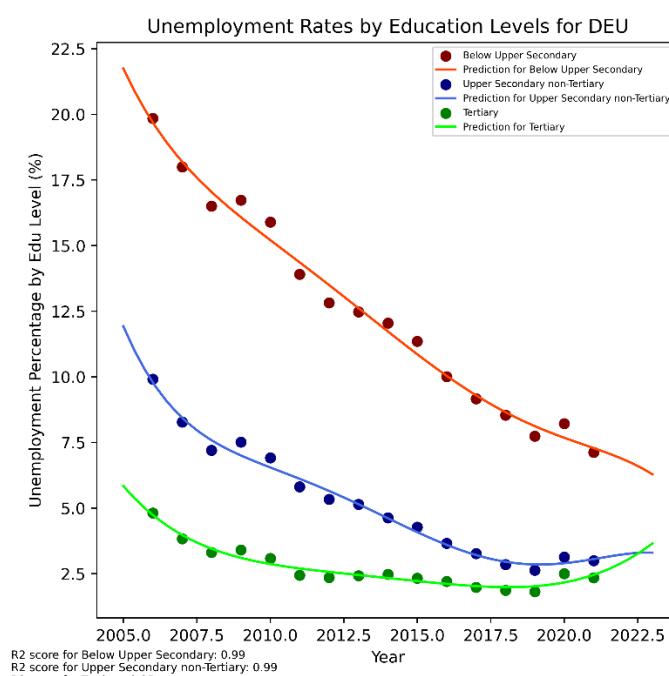
Netherlands is another outlier in this education level.

Tertiary:

cluster	countries
Cluster 0	['AUS', 'CAN', 'SWE', 'CHE', 'TUR']
Cluster 1	['DEU']
Cluster 2	['BEL', 'DNK', 'FIN', 'FRA', 'ITA', 'NLD', 'ESP']
Cluster 3	['JPN', 'USA', 'GBR']

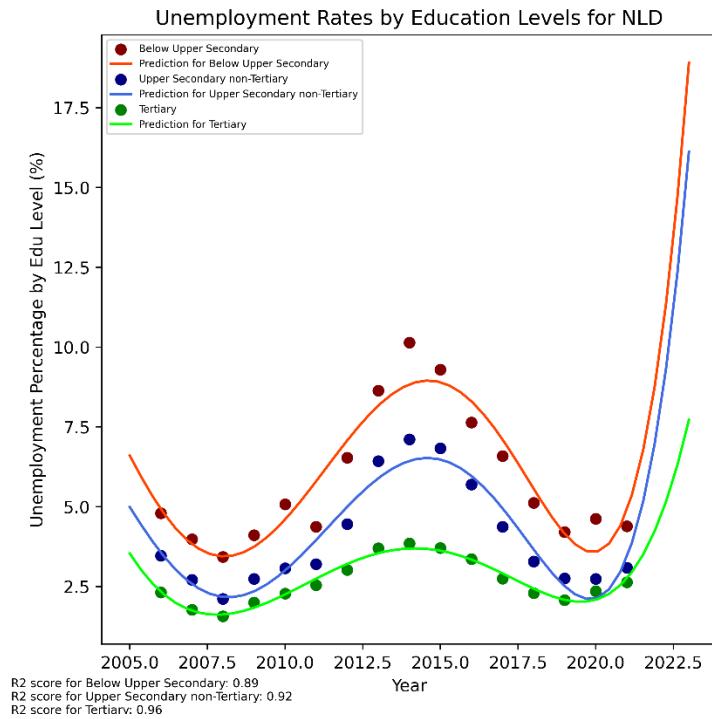
The rate of unemployed high educated people rate is different in Germany.

Germany



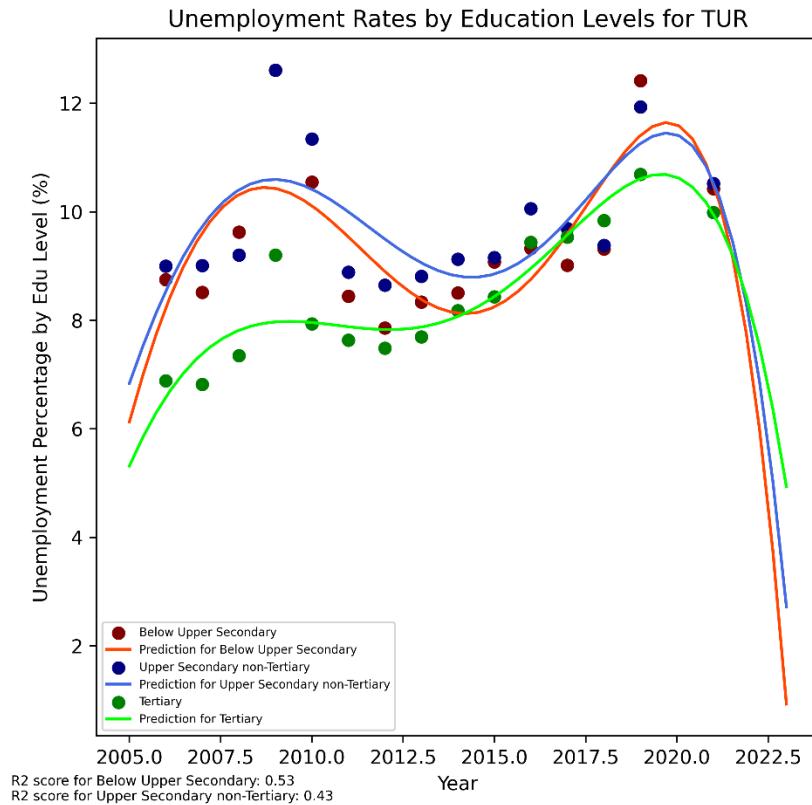
- Germany is the country with the lowest unemployment rate at any level of education.
- We can say that even someone who does not have a tertiary education can easily find a job in Germany today.

Netherlands



- The Netherlands' chart is different from many other countries. While there was an increase in the unemployment rate especially around 2015, we see that this unemployment rate was reduced with the right policy of the government.

Türkiye



- We know that the unemployment rate is increasing in Türkiye. According to education level, this rate is very variable and difficult to predict.

- Unemployment rate according to education in Türkiye is very different from other countries. This ratio is very uncertain at every education level. We see that the unemployment rate of those with tertiary education is sometimes higher than those with a lower education level.
- Although these rates vary a lot, they are not values that are too far from other countries. Therefore, the algorithm did not put Türkiye in a different cluster.

iii. Long Term Unemployment

This dataset is retrieved from “OECD Data” as ".csv" file. This dataset can be reached from <https://data.oecd.org/unemp/long-term-unemployment-rate.htm#indicator-chart> with “Long-term unemployment rate” title.

These are all titles in this dataset:

“LOCATION”, “INDICATOR”, “SUBJECT”, “MEASURE”, “FREQUENCY”, “TIME”, “Value”, “Flag Codes”

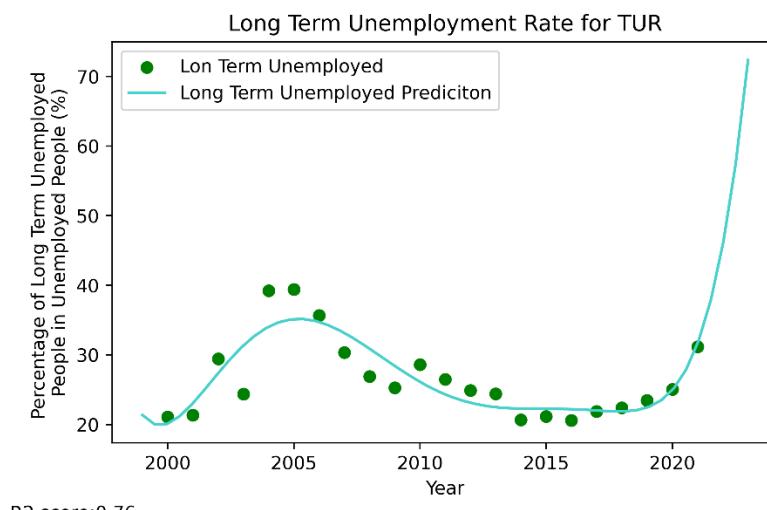
The long-term unemployment rate is the rate of people who have been unemployed for 12 months or more.

The higher rate means unemployment problem is greater in that country.

cluster	countries
Cluster 0	['AUS', 'BEL', 'FIN', 'FRA', 'DEU', 'ITA', 'NLD', 'SWE', 'CHE']
Cluster 1	['TUR']
Cluster 2	['USA']
Cluster 3	['CAN', 'DNK', 'JPN', 'ESP', 'GBR']

We see both Türkiye and USA have different values than other countries.

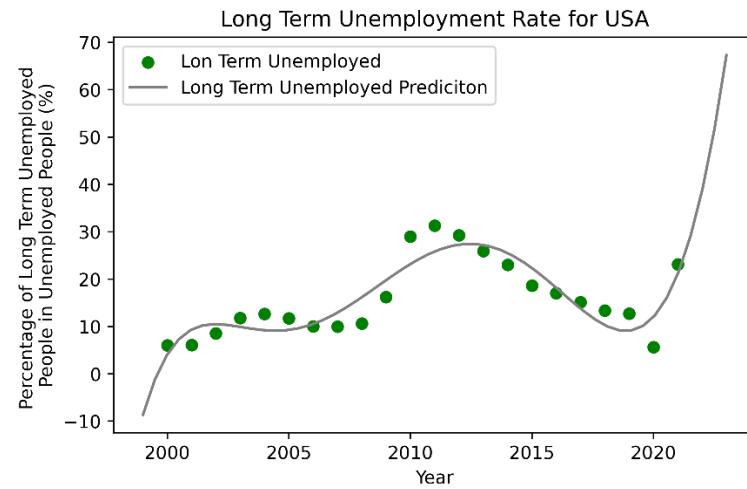
Türkiye



- Although the long-term unemployment problem in Turkey reached 40 percent in 2005, the country solved this problem afterward. After 2018, we see that this rate started to increase again.

- Although this problem sometimes seemed to have been solved in the country, long-term unemployment was worth at least 20 percent which is higher than most of the other countries.

USA



- The long-term unemployment rate is much lower in the USA than in other countries.
- While this rate is 30% at the highest, in some countries the lowest value is 30%.
- The lowest rate is 7% and there are not many countries in the world that reach this rate. Countries that reached these rates in the world could not stay at that level for a long time.

iv. Youth Unemployment

This dataset is retrieved from “OECD Data” as ".csv" file. This dataset can be reached from <https://data.oecd.org/unemp/youth-unemployment-rate.htm#indicator-chart> with “Youth unemployment rate” title.

These are all titles in this dataset:

“LOCATION”, “INDICATOR”, “SUBJECT”, “MEASURE”, “FREQUENCY”, “TIME”, “Value”, “Flag Codes”

Youth unemployment rate is the rate of young people who can work in that country, even though they are looking for a job. Young people need to be able to find work easily when they graduate in countries with strong employment and economy.

Women:

cluster	countries		
Cluster 0	['CAN']		
Cluster 1	['AUS', 'BEL', 'DNK', 'FIN', 'FRA', 'DEU', 'ITA', 'NLD', 'ESP', 'SWE', 'CHE', 'TUR']		
Cluster 2	['JPN', 'USA', 'GBR']		

Canada is only outlier in these clusters.

Men:

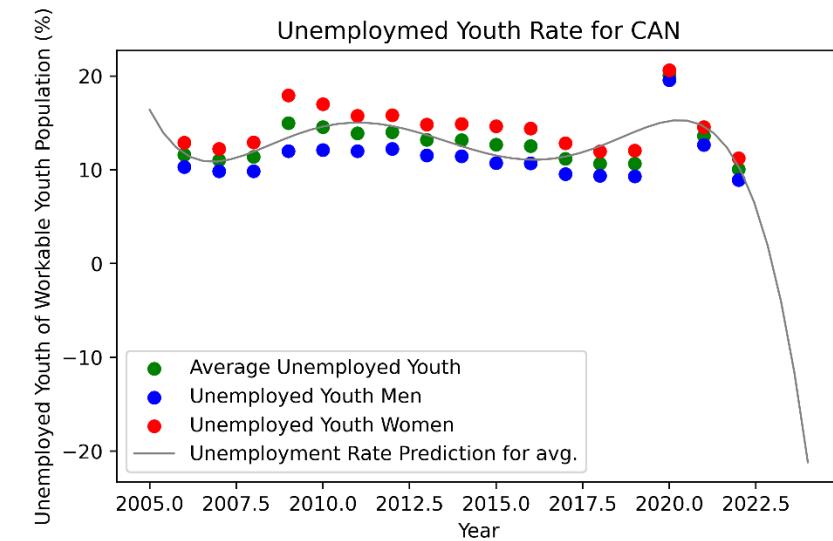
cluster	countries
Cluster 0	['BEL', 'FIN', 'FRA', 'NLD']
Cluster 1	['AUS', 'CAN', 'DEU', 'ITA', 'ESP', 'SWE', 'CHE', 'TUR']
Cluster 2	['DNK', 'JPN', 'USA', 'GBR']

Total:

cluster	countries
Cluster 0	['BEL', 'FIN', 'FRA', 'NLD']
Cluster 1	['AUS', 'CAN', 'DEU', 'ITA', 'ESP', 'SWE', 'CHE', 'TUR']
Cluster 2	['DNK', 'JPN', 'USA', 'GBR']

Clusters are exactly same in Youth Unemployment Rate for men and total.

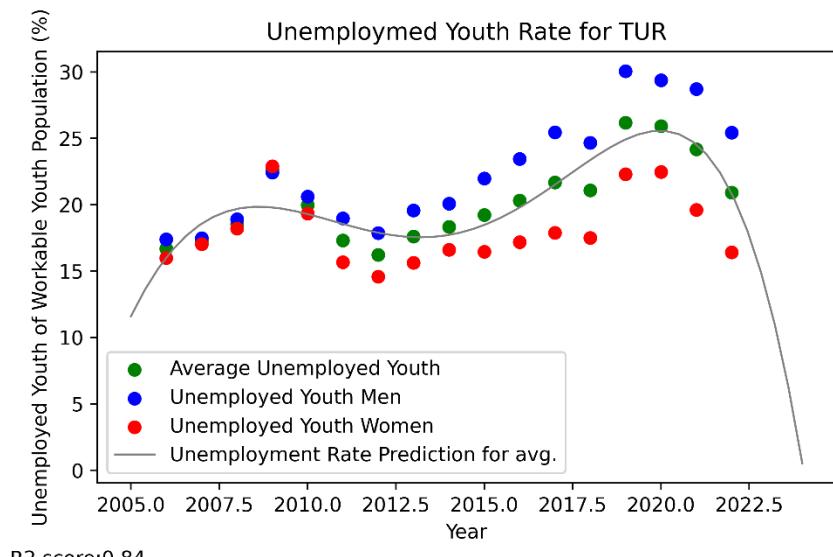
Canada



R2 score: 0.46

- Unemployed youth women rate graph is different than other countries. It is not really different as value, but the graph is more stable than other countries.

Türkiye



R2 score: 0.84

- Youth unemployment rate in Turkey has an increasing graph. Although Turkey is not the only country facing this problem, it does not increase that fast in many countries or it is more stable in many countries.
- Unemployed youth women rate is below Unemployed youth man rate in Türkiye for the last 15 years. This is not very common in other countries. This may be because young women choose to be housewives. Because housewifery is quite common in Turkey.

We examined the unemployment rate from different perspectives in these data. The unemployment rate increases or decreases according to the economic situation and the policies implemented by the governments. The long-term unemployment rate is a data shows the status of the unemployed for 12 months or longer, and the long-term increase in the unemployment rate triggers the long-term unemployment rate. Junankar proved that after a while the unemployment rate increased, the long-term unemployment rate also increased with a lag (2011). For this reason, it is expected that the long-term unemployment rate will increase 12 months after the unemployment rate starts to rise.

The fluctuations in the unemployment rate affect the young people more. The short-term effects of the unemployment rate on the youth unemployment are "30-70%" and the long-term "60-80" times higher (Bayrak, 2018). So, it would not be wrong to say that the youth unemployment rate is much more fragile. While the unemployment rate is expected to decrease as the education level increases, the reality is not like this. Hulsman et al. has expressed the problem of finding jobs for those with higher education, and Europe has focused on implementing the necessary policies (2002). We look forward to seeing the outcome of policies such as the Bologna Process.

It is also clear that the unemployment rate differs by gender. According to Azmat et al. proved that occupations vary according to gender, but they could not find a clear reason in their research (2006).

It should not be forgotten that all these factors are interconnected and can trigger each other. Unemployment is a big problem for countries. With the right policy and economic management, this problem can be solved.

c. Job

i. Job Strain

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Labour" section with "Job strain index" title.

These are all titles in this dataset:

"LOCATION","Country","MEA","Overall measure","VAR","Components","POP","Age","SEX","Sex","EDU","Education","TIME","Time","Unit Code","Unit","PowerCode Code","PowerCode","Reference Period Code","Reference Period","Value","Flag Codes","Flags"

Job Strain can be explained as the difficulty of working hours and working conditions. At the same time, areas such as career opportunities and self-development can be included into this definition.

Job Strain is important for understanding the psychology of the employee. Because people who work under difficult conditions may feel under pressure, get very tired mentally or physically and cannot work efficiently.

We will examine Job Strain under 2 subheadings: High Level of Job Demands, Low Level of Job Resources

1. High Level of Job Demands

High Level of Job Demands can be defined as expecting the employee to do too much work. Long and inflexible hours, life-threatening working conditions are some of them. Long working hours indicates "working more than 60 hours in a week".

Physical Health Risk Factors

cluster	countries
Cluster 0	['AUS', 'CAN', 'JPN', 'NLD']
Cluster 1	['FIN', 'FRA', 'ESP', 'SWE']
Cluster 2	['BEL', 'DNK', 'DEU', 'ITA', 'TUR', 'GBR']
Cluster 3	['USA']

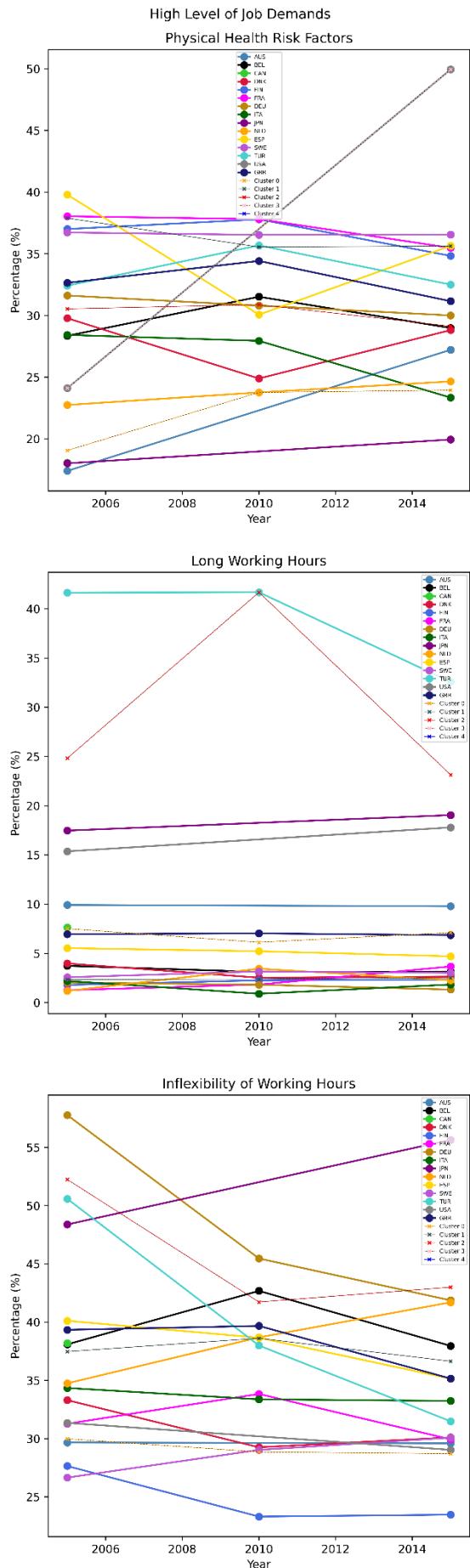
The clusters appear to be almost equally divided into 3 parts except for the USA.

Long Working Hours

cluster	countries
Cluster 0	['AUS', 'CAN', 'ESP', 'GBR']
Cluster 1	['BEL', 'DNK', 'FIN', 'FRA', 'DEU', 'ITA', 'NLD', 'SWE']
Cluster 2	['JPN', 'TUR', 'USA']

Inflexibility of Working Hours

cluster	countries
Cluster 0	['AUS', 'DNK', 'FIN', 'FRA', 'SWE', 'USA']
Cluster 1	['BEL', 'CAN', 'ITA', 'NLD', 'ESP', 'GBR']
Cluster 2	['DEU', 'JPN', 'TUR']



In general, the outstanding data about “High Level of Job Demands” are:

- Jobs with physical health risk factors in the US have increased significantly in 10 years and are by far the top.
- There is also an increase in jobs with physical health risk factors in Australia.
- As we saw in the economic data, people in Turkey earn very little income. Here we see that they work both the cheapest and the longest. Although the rate of working for long hours is decreasing, this rate is still very high.
- Japan and the USA are in the same cluster with Turkey, and they also have a high rate of working for long hours, but not as much as Turkey.
- Türkiye, Germany and Japan are the countries with the least flexibility in working hours.

2. Low Level of Job Resources

Low Level of Job Resources can be explained as career development and advancement opportunities offered by the employer to the employee.

Work Autonomy is to give its employees the opportunity to work in the field that suits them. Learning Opportunities is to provide its employees with the training they need to advance in their field.

Training and Learning is to provide sufficient support to the employee through education such as workshops and seminars to gain experience in the field.

Opportunity for Career Advancement is to provide opportunity and support to its employee for career advancement. Getting promoted at work is an example of this.

Work Autonomy and Learning Opportunities:

cluster	countries
Cluster 0	['AUS', 'BEL', 'CAN', 'FIN', 'FRA', 'DEU', 'USA', 'GBR']
Cluster 1	['DNK', 'NLD', 'SWE']
Cluster 2	['ITA', 'JPN', 'ESP', 'TUR']

Training and Learning:

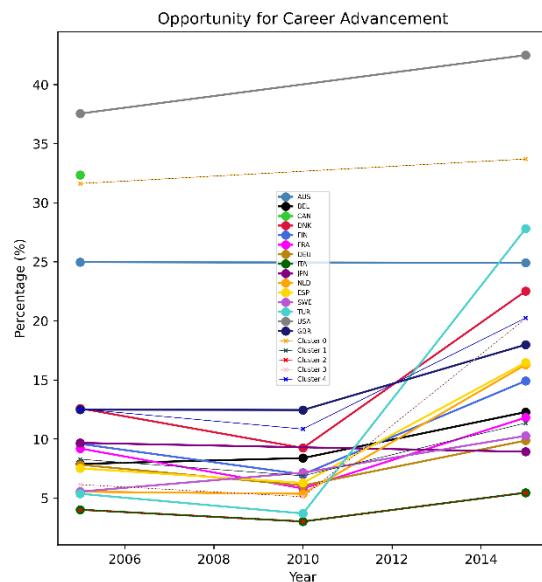
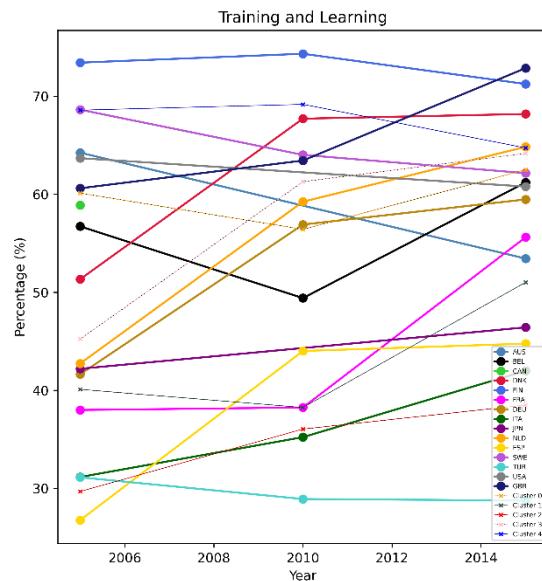
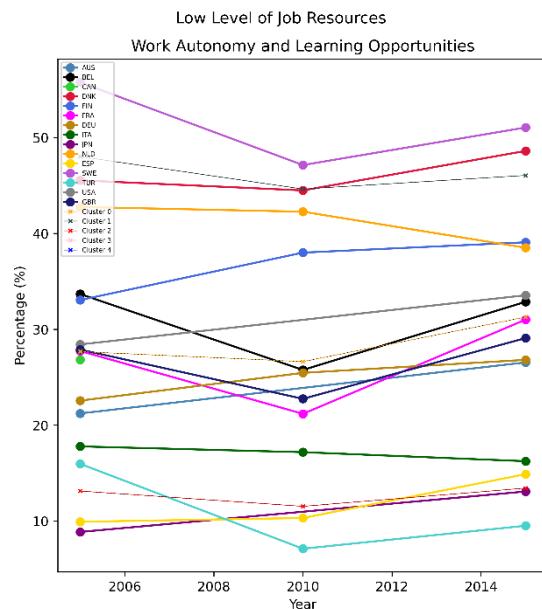
cluster	countries
Cluster 0	['AUS', 'BEL', 'CAN', 'GBR']
Cluster 1	['FRA', 'JPN']
Cluster 2	['ITA', 'ESP', 'TUR']
Cluster 3	['DNK', 'DEU', 'NLD']
Cluster 4	['FIN', 'SWE', 'USA']

Opportunity for Career Advancement:

cluster	countries
Cluster 0	['AUS', 'CAN', 'USA']
Cluster 1	['BEL', 'FIN', 'FRA', 'DEU', 'JPN', 'SWE']
Cluster 2	['ITA']
Cluster 3	['NLD', 'ESP', 'TUR']
Cluster 4	['DNK', 'GBR']

Italy is in Cluster 2 alone among the clusters.

We expect Denmark and Great Britain to have similar values.



In general, the outstanding data about “Low Level of Job Resources” are:

- Denmark, the Netherlands, and Sweden are the countries that give the employee the most opportunity to choose a field and provide support in the field.
- Türkiye, Spain, and Italy are the countries that provide the least training to their employees.
- The countries that support their employees the most to develop their career are Canada, Australia and by far the USA.
- Italy is the country that offers the least career development opportunities to its employees. Although the ratios of Italy and Turkey are close to each other, Turkey rose from second last to second place with its rise in 2014.
- Türkiye is generally below in “Low Level of Job Resources” charts. In other words, Türkiye does not provide sufficient opportunities to its employees and does not support.

The reasons for the job strain are too many job demands and too few job resources. These factors are measured as we explained above. Job quality, workload, and work-life balance are interconnected subjects, and all of them impact health. Working under prolonged stress negatively affects human health. As Picatoste also stated, a better healthcare system will be necessary in the future because working under challenging conditions affects people's health (2017). Building a new healthcare system and providing necessary financial support to prevent people from working long hours are within the governments' hands. Governments should always prioritize people's health above everything else.

ii. Job Quality

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Labour" section with "Job quality by skill" title.

These are all titles in this dataset:

"LOCATION", "Country", "MEA", "Overall measure", "VAR", "Components", "POP", "Age", "SEX", "Sex", "EDU", "Education", "TIME", "Time", "Unit Code", "Unit", "PowerCode Code", "PowerCode", "Reference Period Code", "Reference Period", "Value", "Flag Codes", "Flags"

The quality of the work people do is also very important. Every individual wants to feel safe in his job and to have the conditions of his job suitable. In this data set, we will examine the confidence of the employees in their work and their satisfaction with the environment they work with.

a. Labour Market Insecure

We will examine whether people feel safe in their job, based on their education and ability level.

Low Skilled:

cluster	countries
Cluster 0	['BEL', 'FIN', 'NLD', 'CHE']
Cluster 1	['AUS', 'CAN', 'DNK', 'FRA']
Cluster 2	['JPN']
Cluster 3	['ESP']
Cluster 4	['ITA', 'SWE', 'TUR', 'USA', 'GBR']

Medium Skilled:

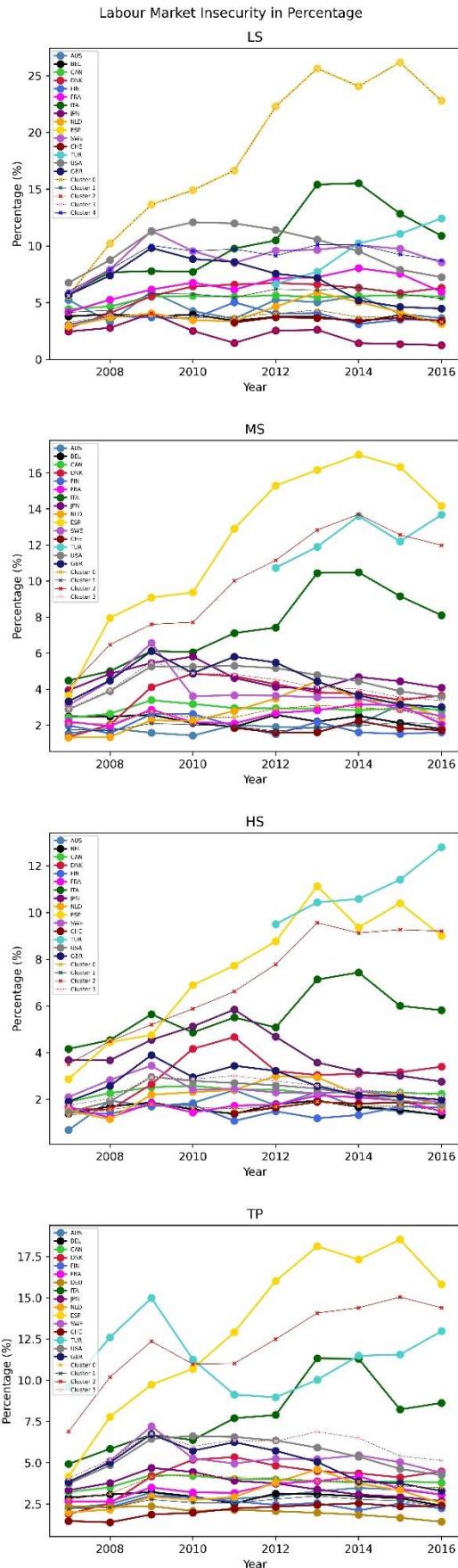
cluster	countries
Cluster 0	['BEL', 'CAN', 'FRA', 'NLD']
Cluster 1	['AUS', 'FIN', 'CHE']
Cluster 2	['ITA', 'ESP', 'TUR']
Cluster 3	['DNK', 'JPN', 'SWE', 'USA', 'GBR']

High Skilled:

cluster	countries
Cluster 0	['AUS', 'BEL', 'FIN', 'FRA', 'CHE']
Cluster 1	['JPN']
Cluster 2	['ITA', 'ESP', 'TUR']
Cluster 3	['CAN', 'DNK', 'NLD', 'SWE', 'USA', 'GBR']

Total:

cluster	countries
Cluster 0	['CAN', 'DNK', 'FRA', 'JPN']
Cluster 1	['AUS', 'BEL', 'FIN', 'DEU', 'NLD', 'CHE']
Cluster 2	['ESP', 'TUR']
Cluster 3	['ITA', 'SWE', 'USA', 'GBR']



In general, the outstanding data are:

- Spain ranks top at all skill levels.
- People in Turkey were ranked last in terms of income at all skill levels. However, they have full confidence in their work, except for the low skilled.
- After Spain and Turkey, Italy is the most confident people in their jobs.
- Although income is highest in Switzerland at every skill level, we see that Swiss people are not very confident in their work.

b. Quality of the Working Environment

Apart from the income and confidence of people in their jobs, the conditions in which they work are also very important. Many environmental factors such as being exposed to mobbing at work and risk of life are effective in this. We will take a closer look at this in this data.

Low Skilled:

cluster	countries
Cluster 0	['CAN', 'FRA', 'DEU', 'JPN', 'ESP', 'SWE']
Cluster 1	['BEL', 'ITA']
Cluster 2	['TUR']
Cluster 3	['AUS', 'DNK', 'FIN', 'NLD', 'USA', 'GBR']

Medium Skilled:

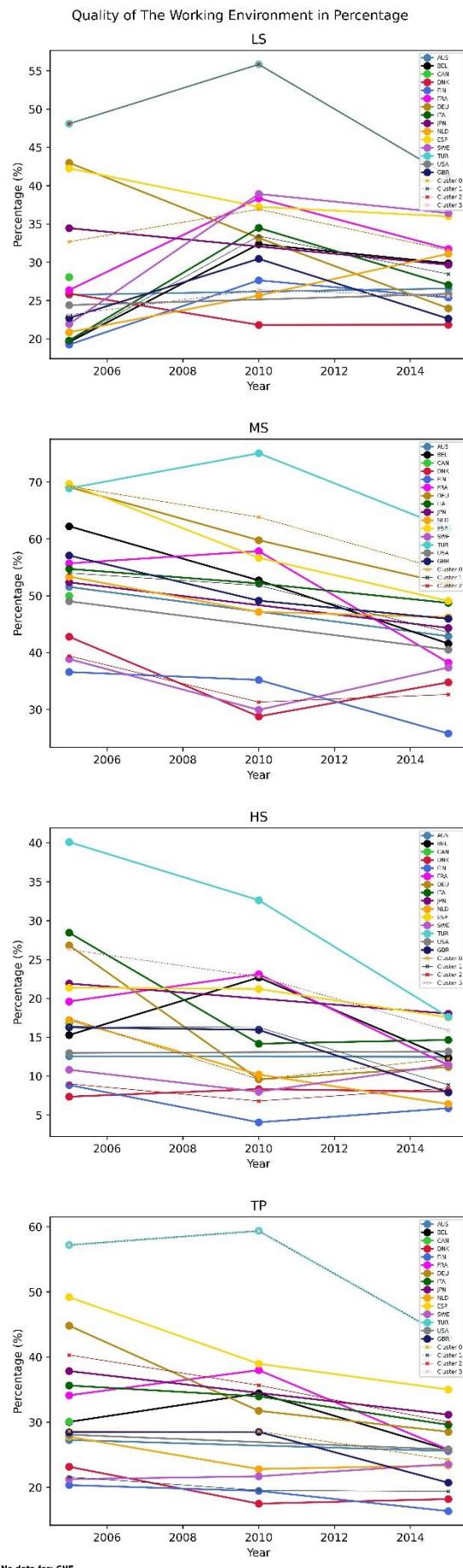
cluster	countries
Cluster 0	['DEU', 'ESP', 'TUR']
Cluster 1	['AUS', 'BEL', 'CAN', 'FRA', 'ITA', 'JPN', 'NLD', 'USA', 'GBR']
Cluster 2	['DNK', 'FIN', 'SWE']

High Skilled:

cluster	countries
Cluster 0	['AUS', 'DEU', 'USA']
Cluster 1	['BEL', 'CAN', 'NLD', 'GBR']
Cluster 2	['DNK', 'FIN', 'SWE']
Cluster 3	['FRA', 'ITA', 'JPN', 'ESP', 'TUR']

Total:

cluster	countries
Cluster 0	['AUS', 'BEL', 'CAN', 'NLD', 'USA', 'GBR']
Cluster 1	['DNK', 'FIN', 'SWE']
Cluster 2	['FRA', 'DEU', 'ITA', 'JPN', 'ESP']
Cluster 3	['TUR']



In general, the outstanding data are:

- Although people in Turkey feel confident with their job, they are not satisfied with their working conditions in every skill level. Because the workload in Turkey is quite high, they work very long hours, and their income is very low.
- Denmark, Finland, and Sweden are in the same cluster at almost every skill level, and when we look at the chart, we see that their values are mostly really close to each other. These countries are usually at the bottom of the chart and people are satisfied with their working conditions.

The satisfaction derived from earned money, trust in the job, and working conditions are subjective factors that vary from person to person. Therefore, whether a job is difficult or easy is a subjective matter that varies from individual to individual. Nonetheless, with this dataset based on surveys, we were able to examine people's opinions about their jobs in the country. Findlay observed in his research that satisfaction, commitment, health, and psychological well-being have an impact on job quality (2013). Many of these factors are related to physical and mental health. Although these factors have an influence on an individual's perception of job quality, it should be remembered that they are subjective.

d. Gross Domestic Spending on R&D

This dataset is retrieved from "OECD Data" as ".csv" file. This dataset can be reached from <https://data.oecd.org/rd/gross-domestic-spending-on-r-d.htm> with "Gross domestic spending on R&D" title.

These are all titles in this dataset:

"LOCATION", "INDICATOR", "SUBJECT", "MEASURE", "FREQUENCY", "TIME", "Value", "Flag Codes"

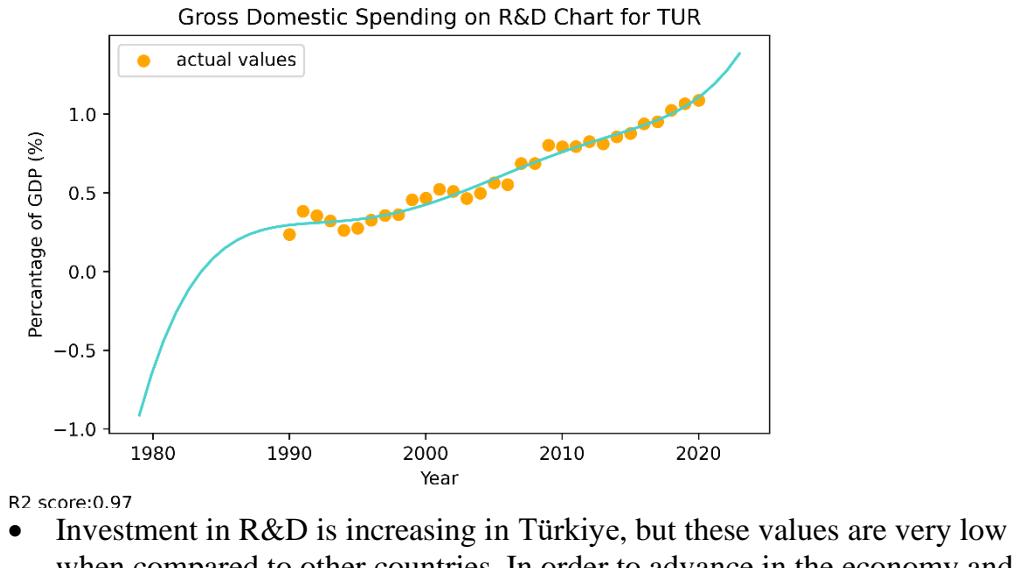
Gross Domestic Spending on R&D helps us see the amount of investment in R&D by companies in that country. If companies want to be successful in their fields, they must give importance to R&D. Therefore, the investment spent on R&D will help us understand how successful companies in that country are.

cluster	countries
Cluster 0	['AUS', 'BEL', 'CAN', 'DNK', 'FRA', 'NLD', 'GBR']
Cluster 1	['ITA', 'ESP']
Cluster 2	['FIN', 'DEU', 'JPN', 'SWE', 'CHE', 'USA']
Cluster 3	['TUR']

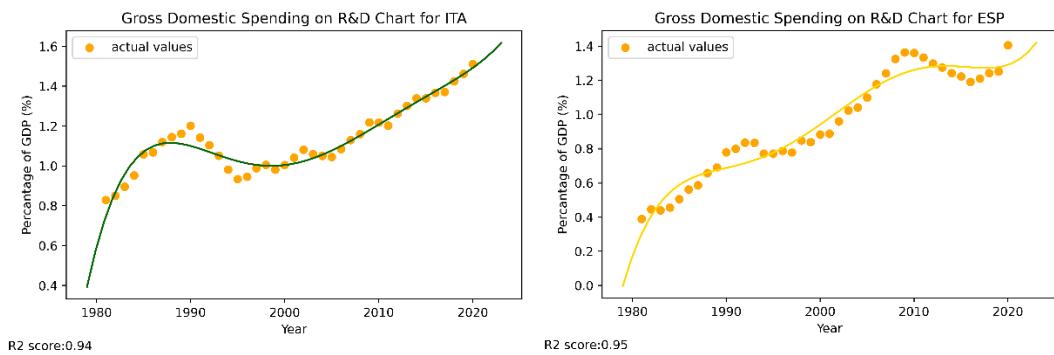
Türkiye is an outlier in these clusters.

Italy and Spain are in same cluster.

Türkiye



Italy and Spain



Governments and companies invest in Research & Development (R&D) for the advancement of a particular sector, science, and technology. Although these investments may appear to be unproductive, they actually contribute significantly to science and technology. Countries with a strong economic status allocate more resources to R&D. On the other hand, countries with weaker economic conditions are unable to allocate a significant amount of resources to R&D as priority is given to vital areas such as health and security. According to Lederman and Maloney's research, investment in R&D leads to an increase in the net export of natural resources and these countries tend to experience faster growth compared to others (2003). Therefore, even if a country is experiencing economic difficulties, investing in R&D to boost natural resource exports could be a move that helps overcome those challenges. This opportunity seems particularly favorable for countries like Turkey at present.

10. Work Life Balance

Work-Life Balance is essential for mental and physical health and allows us to examine the time people devote to their work and personal lives.

A. Work Life Balance

This dataset is retrieved from OECD Stats as ".csv" file. This dataset can be reached from <https://stats.oecd.org/> in the "Education and Training" section with "Work-life Balance" title.

These are all titles in this dataset:

"COUNTRY", "Country", "ISC11A", "ISCED-A 2011", "SEX", "Gender", "AGE", "Age", "PIAAC_CATEGORY", "Category", "INDICATOR", "Indicator", "MEASURE", "Measure", "YEAR", "Reference year", "Value", "Flag Codes", "Flags"

Work Life Balance is how people balance work life, social life, and family life. Work Life Balance can be measured by factors such as whether they can spare time for family from work life, flexibility of working hours, difficulties in concentrating on work due to family problems.

Difficult For Them to Concentrate at Work:

Adults Who Reported That Over the Last 12 Months It Has Been Difficult for Them to Concentrate at Work Because Of Their Family Responsibilities

cluster	countries
Cluster 0	['FRA', 'NLD', 'ESP']
Cluster 1	['TUR']
Cluster 2	['AUS', 'BEL', 'ITA', 'USA', 'GBR']

Difficult For Them to Fulfil Their Family Responsibilities:

Adults Who Reported That Over the Last 12 Months It Has Been Difficult for Them to Fulfil Their Family Responsibilities Because Of The Amount of Time They Spend at Work

cluster	countries
Cluster 0	['AUS', 'BEL', 'FRA', 'ITA', 'USA']
Cluster 1	['NLD']
Cluster 2	['ESP', 'GBR']
Cluster 3	['TUR']

Mean Number of Hours Worked Per Week:

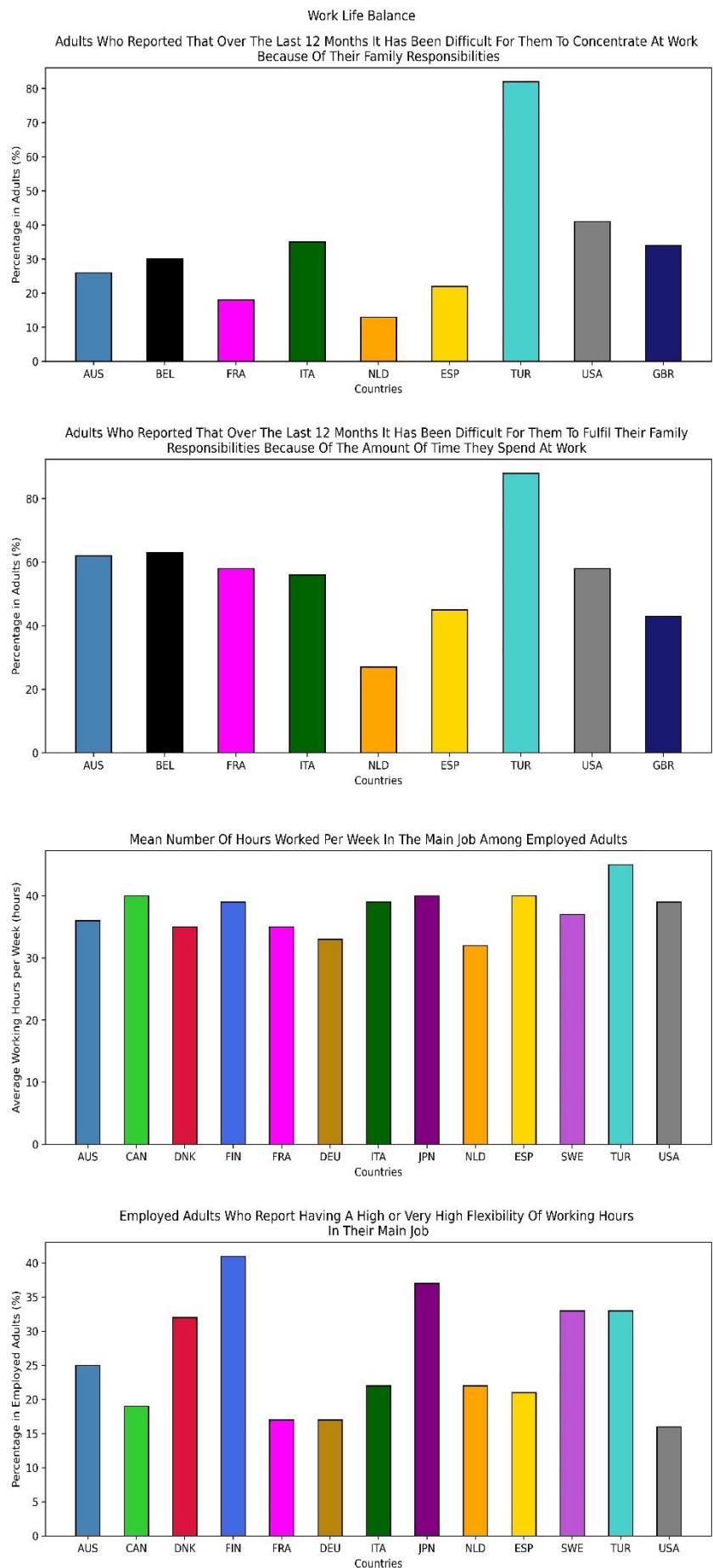
Mean Number of Hours Worked Per Week in The Main Job Among Employed Adults

cluster	countries
Cluster 0	['AUS', 'CAN', 'FIN', 'ITA', 'JPN', 'ESP', 'SWE', 'USA']
Cluster 1	['DNK', 'FRA', 'DEU', 'NLD']
Cluster 2	['TUR']

High or Very High Flexibility of Working Hours:

Employed Adults Who Report Having a High or Very High Flexibility of Working Hours in Their Main Job

cluster	countries
Cluster 0	['CAN', 'FRA', 'DEU', 'USA']
Cluster 1	['AUS', 'ITA', 'NLD', 'ESP']
Cluster 2	['DNK', 'FIN', 'JPN', 'SWE', 'TUR']



In general, the outstanding data are:

- The majority of people in Türkiye cannot concentrate on work due to family responsibilities. this rate is twice that of the USA, which is in the 2nd place.
- Likewise, the majority of people in Turkey cannot fulfill their family responsibilities due to the time they spend at work.
- Most people in the Netherlands have no problem concentrating on their work and fulfilling their family responsibilities.
- Among these countries, Turkey is the country that works the most in a week.
- USA, Canada, France, and Germany do not have much flexibility in working hours, while Denmark, Finland, Sweden, Japan, and Turkey do not have problems with their flexibility in working hours.
- When we look at the data above, people in Turkey cannot fulfill their family responsibilities due to the time they spend at work and cannot concentrate on their work due to these responsibilities.

Employers need to reduce working hours and support employees in this regard so that employees can spend time with their families in Turkey. In this way, the problem of concentrating on the work of the employees is solved and they work more efficiently. In other words, people should work shorter and more efficiently instead of working for long hours and inefficiently.

Work-life balance is very important for the mental health of people. Because human is a social being and needs to be socialized. Even if he can't take time to socialize, he needs to make time for his family. Adults earn money for their family and their own welfare. Trying too hard to earn money and not being able to devote time to their families negatively affects both their personal and family well-being. Having flexible working hours helps the individual to establish a work-life balance. Chung's research findings reveal that flexible work positively affects work-life balance and family functioning. Also, he states that women who engage in flexible work to fulfill domestic responsibilities tend to have their careers negatively impacted, while men prioritize their jobs (2020). Since the roles and priorities of women and men are perceived differently in society, people tend to conform to the expectations of their gender. However, this is yielding to societal judgments. Women and men are equal to each other in the workplace and should be provided with equal opportunities. Individuals should also act according to their individual priorities rather than societal expectations and manage their lives by establishing a work-life balance accordingly.

5. CONCLUSION & IMPACT

Although Switzerland is a country where educated individuals have a lot of influence from politics, the public does not like to participate in politics and the voting rate of people is low. Solijonov says we should be careful about factors that affects only some groups (Solijonov, 2016) and this is quite true. It would not be right to make a definite comment without researching the demographic structure, thinking styles, cultures, and lifestyles of the countries. Many factors can be effective in the non-voting of a certain group.

The people and the government in Turkey do not attach much importance to the environment, Denmark, Finland, and Sweden should set an example to other countries with the importance they give to the environment and the policies they implement.

Also, we should denote that these Policy Stringencies are measured by some complex mathematical calculations and these calculations include tax rates, bans etc. Details on these calculations are in the article by Enrico Botta, Tomasz Koźluk (*Voter Turnout Trends around the World* 2016).

Japan does not pay necessary attention to diet, and as a result, the number of people in poor health goes up.

Although Turkey does not give necessary importance to the environment, they show the necessary importance to nutrition, have enough hospitals and are one of the leading countries in the field of health. Also, there is another article saying that Türkiye is really good in Health field (Gavurova et al., 2021).

While house and rent prices are exorbitantly high, the number of rooms per person is quite low in Türkiye. A large portion of household income in Turkey goes to rent.

The expenditure on rent is not much, and the number of rooms per person is quite adequate in Canada and Australia. In addition, house prices in Italy and Finland do not change much over the years. As some researchers (such as Coşkun) expected, there were problems in housing supply in Turkey and house prices increased exponentially. If there is no solution to this problem in Turkey, it seems that it will continue to increase.

Income inequality in Türkiye and USA is very high, while income inequality in Belgium, Denmark, Finland, and Sweden is very low.

In addition, Turkey's currency is depreciating at an extreme level and since income inequality is high, the gap between the rich and the poor is increasing, while the percentage of the middle-income part is decreasing. Also, Muller revealed that there is a link between education level and income inequality (Muller, *Education, income inequality, and mortality: A multiple regression analysis* 2002). Education is the key to everything, including the economy. There is no success in any field without education. Therefore, it is not possible for a country to develop without a successful education system.

Income is very low regardless of education level, students' scores in PISA are quite low and overqualification is high in Turkey.

Finland and Japan are successful in PISA. Skill mismatch is low in Finland. Rate of foreign students in Great Britain is quite high.

The crime rate is quite high in the USA, the traffic accident rate is high in Türkiye and the USA. According to the study of academicians at Balikesir University, there is a direct proportional relationship between unemployment and crime rate (Dursun et al., 2011). The reason why the crime rate is so high may be the high unemployment rate in Turkey.

People cannot spend a lot of money on socialization in Turkey. As Coşkun predicted, house prices have gone really high in Türkiye (Coskun, The global financial crisis and the Turkish housing market: Is there a success story? 2011). As house prices rose, rents rose. People's income is very low in Turkey and most of this income goes to rent.

The unemployment rate has decreased steadily in Germany and suitable jobs can be found at any level of education. The unemployment rate in Turkey is very high at all levels of education and the values are close to each other. In other words, the education received has no effect on employment. The gap between female and male unemployment rates is narrowing, but still quite large. Long-term unemployment is very high, the rate of young unemployed is increasing. The job strain is too high, and the employees are not provided with sufficient career support in Turkey.

There is no balance between family life and work in Turkey. While this is a problem that makes people mentally and physically tired, it also reduces productivity. For this reason, employers can make their employees more productive by letting them work shorter hours.

All these fields are very important for wellbeing of countries. However, the fact that a country is a leader in a field does not mean that it is a livable or successful country. For this reason, only countries that are successful in many of these areas makes their citizens live in welfare.

6. DISCUSSION & FUTURE WORK

Since it is the only research that has been done in such a wide area using data mining, there is no other research to compare. This research has been prepared to explain how data mining methods can be used in many fields and how helpful the results are. Wellbeing of countries is a very open-ended topic, and more data can be found every time. We tried to find successful countries and unsuccessful countries at least 1 dataset in every field. We used 36 datasets in total to find answers. We made sure that each of these data was from a reliable source and only used data from the OECD. Data is available from many sources other than the OECD, but we cannot be sure of its reliability.

A more detailed study can be done by finding more datasets as future studies. But too much data is not always good either. Also, care should be taken to get the data from a reliable source. Research can be developed by creating clusters again by looking at the number of clusters that the obtained clusters have in common. Alternatively, a new similarity matrix can be constructed with the average or weighted average of all similarity matrices in a field (e.g., Health). Thus, for each area, the generally good and bad countries are separated.

7. REFERENCES

- [1] Alsan, M., & Goldin, C. (2019). Watersheds in child mortality: The role of effective water and sewerage infrastructure, 1880–1920. *Journal of Political Economy*, 127(2), 586-638.
- [2] Atıktürk Taşdelen, D. (2019). *APPLICATION OF DATA MINING IN MEDICAL SCIENCE AND HEALTH CARE* (Publication No. 572189) [Master thesis, Ankara University]. YÖK Ulusal Tez Merkezi.
- [3] A. Grianti, “Cosine similarity matrix using broadcasting in Python,” Medium, <https://towardsdatascience.com/cosine-similarity-matrix-using-broadcasting-in-python-2b1998ab3ff3> (accessed May 30, 2023).
- [4] A. Muller, “Education, income inequality, and mortality: A multiple regression analysis,” The BMJ, <https://www.bmj.com/content/324/7328/23.short> (accessed May 30, 2023).
- [5] A. Solijonov, IDEA, Stockholm, Strömsborg, rep., 2016
- [6] Adema, W., Fron, P., & Ladaïque, M. (2011). Is the European welfare state really more expensive?: indicators on social spending, 1980-2012; and a manual to the OECD social expenditure database (SOCX).
- [7] *Affinity propagation*. (2021, September 11). Wikipedia.
https://en.wikipedia.org/wiki/Affinity_propagation
- [8] Azmat, G., Güell, M., & Manning, A. (2006). Gender gaps in unemployment rates in OECD countries. *Journal of Labor Economics*, 24(1), 1-37.
- [9] B. Gavurova, K. Kocisova, and J. Sopko, “Health System Efficiency in OECD countries: Dynamic Network DEA approach - Health Economics Review,” BioMed Central, <https://healtheconomicsreview.biomedcentral.com/articles/10.1186/s13561-021-00337-9> (accessed May 30, 2023).
- [10]
- [11] Baum, J. I., Kim, I. Y., & Wolfe, R. R. (2016). Protein consumption and the elderly: what is the optimal level of intake?. *Nutrients*, 8(6), 359.
- [12] Bayrak, R., & Tatlı, H. (2018). The determinants of youth unemployment: A panel data analysis of OECD countries. *The European journal of comparative economics*, 15(2), 231-248.
- [13] Blais, A. (2006). What affects voter turnout?. *Annu. Rev. Polit. Sci.*, 9, 111-125.
- [14] Blais, A. (2014). Why is turnout so low in Switzerland? Comparing the attitudes of Swiss and German citizens towards electoral democracy. *Swiss Political Science Review*, 20(4), 520-528.
- [15] Boffetta, P., & Parkin, D. M. (1994). Cancer in developing countries. *CA: a cancer journal for clinicians*, 44(2), 81-90.

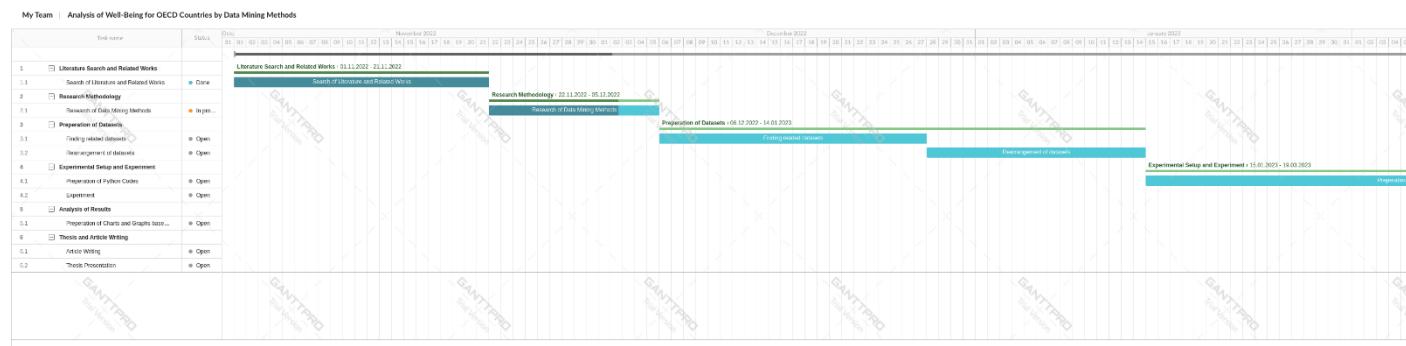
- [16] Britannica, T. Editors of Encyclopaedia (2011, February 3). social welfare program. Encyclopedia Britannica. Retrieved November 27, 2022, from <https://www.britannica.com/topic/social-welfare-program>
- [17] Cambridge Dictionary. (n.d.). Well-being. In Cambridge University's Dictionary. Retrieved November 29, 2022, from <https://dictionary.cambridge.org/dictionary/english/well-being>
- [18] Chernyshenko, O. S., Kankaraš, M., & Drasgow, F. (2018). Social and emotional skills for student success and well-being: Conceptual framework for the OECD study on social and emotional skills.
- [19] Chung, H., & Van der Lippe, T. (2020). Flexible working, work–life balance, and gender equality: Introduction. *Social Indicators Research*, 151(2), 365-381.
- [20] Coşar, K. (2019). ANALYSIS OF OECD HEALTH DATA WITH DATA MINING METHODS (Publication No. 638039) [Master thesis, Marmara University]. YÖK Ulusal Tez Merkezi.
- [21] E. Botta and T. Koźluk, "Measuring environmental policy stringency in OECD countries," *OECD Economics Department Working Papers*, 2014.
doi:10.1787/5jxrjnc45gvg-en
- [22] Findlay, P., Kalleberg, A. L., & Warhurst, C. (2013). The challenge of job quality. *Human Relations*, 66(4), 441-451.
- [23] GeeksforGeeks. (n.d.). *Data Mining Techniques*. GeeksforGeeks. <https://www.geeksforgeeks.org/data-mining-techniques/>
- [24] Georgia Tech Boot Camps. (2021, July 21). *10 Key Data Mining Techniques and How Businesses Use Them*. Georgia Tech Boot Camps. <https://bootcamp.pe.gatech.edu/blog/10-key-data-mining-techniques-and-how-businesses-use-them/>
- [25] Glaeser, E. L., Gyourko, J., & Saks, R. E. (2005). Why have housing prices gone up?. *American Economic Review*, 95(2), 329-333.
- [26] Hassan, M., Nassar, R., & Liu, C. (2014). Effect of government deficit spending on the GDP in the United States. *Journal of Economic and Economic Education Research*, 15(3), 103-112.
- [27] Hoxhaj, Jona & Hysa, Eglantina. (2015). Well-being in Developing Countries.
- [28] Huisman, J., & Kaiser, F. (2002). A Comparative View On Policy Trends In Western European Higher Education. *German Policy Studies/Politikfeldanalyse*, 2(3).
- [29] Immervoll, H., & Scarpetta, S. (2012). Activation and employment support policies in OECD countries. An overview of current approaches. *IZA Journal of Labor Policy*, 1, 1-20.
- [30] Johnstone, N., Haščić, I., Poirier, J., Hemar, M., & Michel, C. (2012). Environmental policy stringency and technological innovation: evidence from survey data and patent counts. *Applied Economics*, 44(17), 2157-2170.
- [31] Junankar, P. N. (2011). The global economic crisis: Long-term unemployment in the OECD.
- [32] Koc, N., & Celik, B. (2015). The impact of number of students per teacher on student achievement. *Procedia-Social and Behavioral Sciences*, 177, 65-70.

- [33] Lederman, D., & Maloney, W. F. (2003). R&D and development. Available at SSRN 402480.
- [34] Lochner, L. (2020). Education and crime. In *The economics of education* (pp. 109-117). Academic Press.
- [35] Manacorda, M., & Petrongolo, B. (1999). Skill mismatch and unemployment in oecd countries marco manacorda. *Economica*, 66(262), 181-207.
- [36] *Measuring well-being and progress: Well-being research*. OECD. (n.d.). Retrieved November 29, 2022, from <https://www.oecd.org/wise/measuring-well-being-and-progress.htm>
- [37] Nadanovsky, P., & Cunha-Cruz, J. (2009). The relative contribution of income inequality and imprisonment to the variation in homicide rates among developed (OECD), South and Central American countries. *Social Science & Medicine*, 69(9), 1343-1350.
- [38] Osberg, Lars & Sharpe, Andrew. (2011). Moving from a GDP-Based to a Well-Being Based Metric of Economic Performance and Social Progress: Results from the Index of Economic Well-Being for OECD Countries, 1980-2009.
- [39] Osberg, Lars & Sharpe, Andrew. (2010). The Index of Economic Well-Being. *Challenge*. 53. 25-42. 10.2753/0577-5132530402.
- [40] Picatoste, X., Aceleanu, M. I., & Şerban, A. C. (2021). Job quality and well-being in OECD countries. *Technological and Economic Development of Economy*, 27(3), 681-703.
- [41] “Pyplot ,” Pyplot tutorial - Matplotlib 3.7.1 documentation, <https://matplotlib.org/stable/tutorials/introductory/pyplot.html> (accessed May 30, 2023).
- [42] Ravallion, M. (2014). Income inequality in the developing world. *Science*, 344(6186), 851-855.
- [43] Rivadeneira, Fabricio & Figueiredo, Adelaide & Figueiredo, Fernanda & Carvajal, Silvia & Rivadeneira, Andrés. (2016). ANALYSIS OF WELL-BEING IN OECD COUNTRIES THROUGH STATIS METHODOLOGY. *HOLOS*. 7. 335. 10.15628/holos.2016.5003.
- [44] Roser, M., Ortiz-Ospina, E., & Ritchie, H. (2013). Life expectancy. *Our world in data*.
- [45] S. Dursun, S. Aytaç, and F. Topbaş, “The effects of unemployment and income on crime: A panel data analysis on Turkey,” DSpace Ana Sayfası, <https://dspace.balikesir.edu.tr/xmlui/handle/20.500.12462/4777> (accessed May 30, 2023).
- [46] Schlesinger, R. B. (2007). The health impact of common inorganic components of fine particulate matter (PM2. 5) in ambient air: a critical review. *Inhalation toxicology*, 19(10), 811-832.
- [47] Sklearn AffinityPropagation, <https://scikit-learn.org/stable/modules/generated/sklearn.cluster.AffinityPropagation.html> (accessed May 30, 2023).
- [48] Solari, C. D., & Mare, R. D. (2012). Housing crowding effects on children’s wellbeing. *Social science research*, 41(2), 464-476.

- [49] Spilimbergo, A. (2009). Democracy and foreign education. *American economic review*, 99(1), 528-543.
- [50] Stanhope, K. L. (2016). Sugar consumption, metabolic disease and obesity: The state of the controversy. *Critical reviews in clinical laboratory sciences*, 53(1), 52-67.
- [51] Tamborini, C. R., Kim, C., & Sakamoto, A. (2015). Education and lifetime earnings in the United States. *Demography*, 52(4), 1383-1407.
- [52] Terry, P., Terry, J. B., & Wolk, A. (2001). Fruit and vegetable consumption in the prevention of cancer: an update. *Journal of Internal Medicine*, 250(4), 280-290.
- [53] Van Noord, J., Kuppens, T., Spruyt, B., & Spears, R. (2022). When and Why People Prefer Higher Educated Politicians: Ingroup Bias, Deference, and Resistance. *Personality and Social Psychology Bulletin*, 49(4), 585–599.
<https://doi.org/10.1177/01461672221077794>
- [54] Varabyova, Y., & Schreyögg, J. (2013). International comparisons of the technical efficiency of the hospital sector: panel data analysis of OECD countries using parametric and non-parametric approaches. *Health policy*, 112(1-2), 70-79.
- [55] Wegman, F., Allsop, R., Antoniou, C., Bergel-Hayat, R., Elvik, R., Lassarre, S., ... & Wijnen, W. (2017). How did the economic recession (2008–2010) influence traffic fatalities in OECD-countries?. *Accident Analysis & Prevention*, 102, 51-59.
- [56] Wikipedia Contributors. (2019, March 3). *Cosine similarity*. Wikipedia; Wikimedia Foundation. https://en.wikipedia.org/wiki/Cosine_similarity
- [57] Wikipedia Contributors. (2019, November 5). *Polynomial regression*. Wikipedia; Wikimedia Foundation.
https://en.wikipedia.org/wiki/Polynomial_regression
- [58] Y. Coskun, “The global financial crisis and the Turkish housing market: Is there a success story?,” SSRN,
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1823304 (accessed May 30, 2023).
- [59] Talend. (n.d.). *16 Data Mining Techniques: The Complete List*. Talend - a Leader in Data Integration & Data Integrity. <https://www.talend.com/resources/data-mining-techniques/>

8. APPENDIX

A. APPENDIX A: GANTT CHART



You can access the high-resolution image of the Gantt Chart via this link:

<https://drive.google.com/file/d/1B2cHBqWZyhjnz9G48nu4PKgheMcAc7DP/view?usp=sharing>

B. APPENDIX B: ALL MATERIALS

You can reach all source codes and output files from link in below:

[OneDrive Link for Access](#)