MASTER THESIS

MSc in Innovation, Knowledge and Economic Dynamics (MIKE-E)

Spring 2019

AALBORG UNIVERSITET

Written by: Karolina Grodzińska

Study Number: 20177027

Supervisor: Roman Jurowetzki

Table of content:

Table of figures:

Abstract

The purpose of this thesis is to discover what are characteristics of innovative economies and what are the factors that make some countries more innovative, through examining the data from Global Innovation Index, one of the most popular rankings that shows the innovative performance of different economies. This research was achieved by explaining the primary theories that relate to innovation, both as a concept and from a systemic point of view, introducing National Systems of Innovation. Additionally, approaches to innovation measurement and competitiveness will be introduced that have been useful in the analysis of differences between countries, and defining limitations of this thesis and used data. The quantitative analysis has been performed through unsupervised machine learning, particularly hierarchical clustering. Both discussion and conclusion focus on limitations related to analyzing aggregate indicators and criticism towards measures of innovativeness used in Global Innovation Index, as it is not possible to univocally state what makes some countries more innovative because the main framework is focused on many phenomena that are not directly related to innovation. However, the ranking can be used to gain many interesting insights on a macroeconomic scale, benchmark different countries based on a number of criteria, and identify advantages or obstacles to innovation.

Keywords: national system of innovation, innovation measurement, competitiveness, aggregate indicators

1. Introduction

Innovation is said to be essential for improvements in living standard and it can affect individuals, institutions, entire economic sectors, and countries in multiple ways. It is also considered a key factor for driving economic progress and competitiveness, for both developed, and developing economies (Oslo Manual, 2018). Therefore, it became important for decision-makers and analysts to understand what are the factors that influence innovation and ways to promote innovation activities. The purpose of this project is to understand what are characteristics of innovative economies and what are the factors that make some countries more innovative. The analysis will be conducted on a macroeconomic level, comparing the performance of different countries in a Global Innovation Index for the year 2018.

The Global Innovation Index (GII) is described as a leading reference on innovation, being the most known ranking for measuring innovative performance on a country level. The Global Innovation Index project was launched in 2007 with the main goal of determining how to find metrics and approaches that better capture the richness of innovation in society and to go beyond traditional measures of innovation, such as the number of research articles or R&D expenditures (Global Innovation Index, 2018).

A key challenge, listed in the Global Innovation Index, is finding metrics that capture innovation as it actually happens in the modern world. Direct official measures, which quantify innovation outputs are considered to remain extremely limited. As a given example, there are no official statistics on the amount of innovative activity measured by the number of new products, processes, or other innovations for any given actor or, even more importantly, any country (Global Innovation Index, 2018).

The Global Innovation Index report addresses some of the challenges related to finding suitable innovation metrics. Some of the key questions raised in the preface include: *How should one better measure innovation and intangible assets in the services sector? How can linkages between innovation actors be better quantified and assessed? How can the more open nature of innovation processes be captured?* (Global Innovation Index, 2018).

Most studies of innovation systems consider mainly generation of innovations, and only to a limited extent with their diffusion, looking only at the supply side of innovation. In order to understand how to succeed in generating economic growth, it is needed to include the demand side as well, including entrepreneurial activity, and business formation (Carlsson, 2006). Moreover, indicators for innovation should not only capture material inputs such as R&D expenditures or human capital inputs, based on the number of years of education. They also need to capture how those resources are used and whether or not the environment promotes further development of knowledge and skills (Lundvall, 2017). Global Innovation Index is trying to address these issues as well, as its framework is divided into innovation inputs and outputs.

Synthetic indicators, like the Global Innovation Index, take into account various aspects that form the technological capability of a country and aggregate them into a single figure. Those indicators are typically macroeconomic, aiming to show the comparison between the positions of different countries and their changes over time. Their main advantage is providing a clear and immediate view on a country's ranking, while the biggest drawback is sacrificing the inherent complexity of the knowledge production and distribution (Archibugi et al., 2009).

There are a few reasons to collect data on national technological capabilities and create innovation indicators. The most important purposes include the use as a source of information for forming public policies, as input for forming firms' strategies, and purely theoretical analysis. Innovation indicators may be used to increase and broaden knowledge of technological change, improve society’s understanding of innovation and its socio-economic effects, or to test theories by researchers. Academics have a strong interest in researching predictive and causal interpretations of innovation outcomes. Managers use innovation studies to gain a deeper understanding of technological advance and to form firms' strategies, especially in a period of fierce competition on a local and international level. Availability of data on the technological capabilities of different countries grants a better understanding of the geographical contexts in which firms can establish and develop their innovative activities, and to benchmark their organization's innovation activities and outcomes. For policymakers, it is important to locate country's position in the global landscape to be able to identify national strengths and weaknesses, to spot technological opportunities, and to assess the effectiveness of the adopted policies. Many governments place innovation at the center of their growth strategies. It is said that reading and interpreting statistics of technological change also provides a fundamental source of information to design and carry out effective innovation policies (Archibugi et al., 2009)(Oslo Manual, 2018).

Therefore, understanding innovation and trying to measure it properly is important for different groups of users. The goal of this thesis is to discover what are the characteristics of the most innovative economies and address some issues related to both innovation measurement and using aggregate indicators.

This assignment is structured as follows. Firstly, there will be a problem formulation and a research question, together with a description why is it important to study competition based on an innovativeness level between different countries. Chapters 3 and 4 concern methodology and data, respectively. In the 5th chapter various theories will be introduced, in order to gain a better understanding of key concepts before the analysis. Chapter number 6 includes analysis, based on findings from the Global Innovation Index, theories, and quantitative study of the data. Sections 7 and 8 present discussion and conclusion of this thesis.

2. Problem formulation

There has been growing interest concerning the competitiveness of countries. Researchers have been trying to explore why do some economies grow much faster than other ones and what are the crucial factors standing behind such differences (Fagerberg et al., 2007).

Within economic and social theories, there is a consensus about the fact that technological change represents the engine of development and progress. To be more specific, innovation is considered to be a determinant of economic growth, productivity, competitiveness, and employment. Therefore, there is a need for finding appropriate measurement tools in order to test and quantify these hypotheses (Archibugi et al., 2009).

The most innovative economies are said to be those that construct appropriate capabilities. However, capabilities tend to be localized and nationally differentiated. Therefore, there can be many successful players in the competitive game, each of them to some extent learning from and interacting with the somewhat alternative paths to capability creation taken by other countries. Following this approach, the pursuit of competitiveness through innovation has become an increasingly important objective of national policies, especially since the role of innovation has risen in the modern knowledge-driven economy, in particular for countries that start behind and wish to catch up (Cantwell, 2005).

2.1 Research question

Based on the problem formulation above, the following research question was formulated for this thesis:

What are the characteristics of innovative economies and what are the factors that make some countries more innovative?

The research question is based on the analysis of one of the most popular rankings, showing the innovative performance of different economies, taking into consideration theoretical wonderings about how innovation should be measured and what are the limitations of such indicators related to innovation. The characteristics will be discovered through analysis of the Global Innovation Index report, combined with theories about national systems of innovation and competitiveness on a macroeconomic scale. The second part of analysis will include quantitative analysis of numerical data that Global Innovation Index is based upon.

3. Methodology

In this thesis, different methods of unsupervised machine learning will be used to identify patterns between observations (countries). The main parts of quantitative analysis will include Principal Component Analysis and clustering.

Principal Component Analysis is based on computing the low-dimensional representation that maximizes data variance. The low-dimensional representation of the data returned by PCA is related to the input feature vectors by a linear transformation. PCA is a type of linear transformation that fits the given dataset into a new coordinate system, in a way where the most significant variance is found on the first coordinate, and each subsequent coordinate is orthogonal (uncorrelated) to the previous one, and explains less of a variance. Besides linearity assumption, PCA also assumes that principal components with larger associated variances represent the most interesting structure, while those with lower variances represent noise (Hayden, 2018)(Yan et al., 2007).

Principal Component Analysis is described as a way of transforming a set of x correlated variables over y samples into a set of p uncorrelated principal components over the same samples. When more principal components are added, more of the original dataset is summarized. Adding additional components makes estimates of the total dataset more accurate but also more unwieldy. In more simple terms - PCA transforms dataset with many variables and simplifies it by changing the original variables into a smaller number of "principal components", which represent the underlying structure of the data. Thus, principal components show the directions in which there is the most variance and where the data is most spread out. This means trying to find a straight line that best spreads the data out when it is projected along it. The first principal component is just this straight line, which shows the most substantial variance in the data (Hayden, 2018).

PCA in this case will be used both as a part of Exploratory Data Analysis to gain some more insights into the dataset and as a part of pre-processing for clustering.

Clustering is described as the unsupervised classification of patterns (observations, data items, or feature vectors) into groups, called clusters. Clustering problem has been addressed in many disciplines, which reflects its broad use as a part of exploratory data analysis. It is useful in decision-making, several exploratory pattern-analysis, grouping, or machine-learning methods, including data mining, document retrieval, image segmentation, or pattern classification (Jain et al., 1999).

Cluster analysis is the organization of a collection of unlabeled patterns, which are usually represented as a vector of measurements, or points in a multidimensional space, into meaningful clusters, based on similarity. Intuitively, patterns grouped together into a valid cluster are more similar to each other than objects belonging to a different cluster. The goal of these algorithms is to create clusters that are coherent internally, but clearly different from one another externally. Thus, objects within a cluster should be as similar as possible to each other, and the patterns from one cluster should be as dissimilar as possible from entities in another (Jain et al., 1999)(Pathak, 2018).

In the case of unsupervised learning techniques, labels are associated with clusters but are also data-driven, meaning that they are obtained solely from the data In many cases, there is little prior information available about the data, and the decision-maker should make as few assumptions about the data as possible. Under these restrictions, clustering methodology is appropriate in particular for the exploration of interrelationships among data points to make a preliminary assessment of their structure (Jain et al., 1999)(Pathak, 2018).

In a broad distinction, there are two ways of clustering data: agglomerative and divisive. An agglomerative approach begins with each observation being clustered into a distinct (singleton) group, and the clusters are successively merged together until a stopping criterion is satisfied. A divisive approach begins with all patterns grouped into a single cluster, and later on being split until a stopping criterion is met (Pathak, 2018).

Clustering algorithms can also be categorized based on how they form clusters. The first approach is called centroid-based clustering, where clusters are represented by a central vector or a centroid. This centroid does not necessarily have to be a member of the dataset. The notion of similarity in this type of clustering is derived by how close a data point is to the centroid of the cluster. It is an iterative clustering method (Jaiswal, 2018).

Another method is called distribution-based clustering and it is closely related to distributional modeling from statistics. This type of clustering is based on the notion of how likely is it for a data point to belong to a certain distribution, for instance, the Gaussian distribution. Thus, data points clustered together belong to the same distribution (Jaiswal, 2018).

Next approach is called density-based clustering, which is based on the search for areas of varied density of data points. Clusters are defined as areas of higher density within the data space, compared to other regions, and the data points in the sparse areas are considered to be noise and/or border points (Jaiswal, 2018).

Another approach is called connectivity-based clustering, also commonly described as "hierarchical clustering". The main idea behind this method is that data points, which are closer together are more similar to each other than to data points further away. Clusters are formed by connecting data points based on their distance, it is started with each data point belonging to individual cluster, and then the closest pairs of points are merged together. Different clusters form at different distances and they can be represented graphically using a dendrogram. These methods do not produce an unique partitioning of the dataset, but rather a hierarchy from which the analyst needs to choose appropriate clusters. Hierarchical clustering is not very robust towards outliers, which may show up as additional clusters or cause other clusters to merge (Jaiswal, 2018)(Pathak, 2018).

Hierarchical clustering is the method of choice for this thesis. Additionally, in order to gain more insights about characteristics of different countries and their importance for innovation, the analysis will be complemented with the use of different theories, and insights from the Global Innovation Index report.

4. Data

The initial cross-sectional data was obtained from the official website of Global Innovation Index, representing the state for the year 2018. From the description in the report, it appears that data was collected for 126 countries, which are supposed to represent 90.8% of the world’s population and 96.3% of global GDP. The innovative performance of each country is measured by 80 different indicators.

Original dataset consisted of 638 observations of 114 variables and even from those numbers it is clear that it needed a lot of preprocessing and transposition of the whole data frame, as the observations were stored in columns and variables in rows, making it more difficult to conduct further analysis. It is said that since most “real world” datasets are not tidy, around 80% of data analysis is spent just on the process of cleaning and preparing the data. The preparation is not only the first step before actual analysis, but must be repeated multiple times over the course of analyzing the data, as new problems arise or new observations are added (Wickham, 2014).

Transposing the original dataset was the first and most important step of preprocessing. Additionally, both columns and rows needed to be renamed to make it easier for further analysis. In the final dataset each observation should be named after the country it stores scores for, and each column involves the name of the variable, representing scores for different indicators measured in Global Innovation Index. From the number of observations and variables, it also seemed that some countries appeared in the dataset a few times. Some columns did not include any data. Therefore, only the rows that included actual observations were kept in the further analysis, ensuring that each country appeared only once in the final dataset so that it would actually contain 126 countries as observations. Moreover, the type of data needed to be changed, the variables in the original dataset were read as character string type and needed to be changed to numerical values.

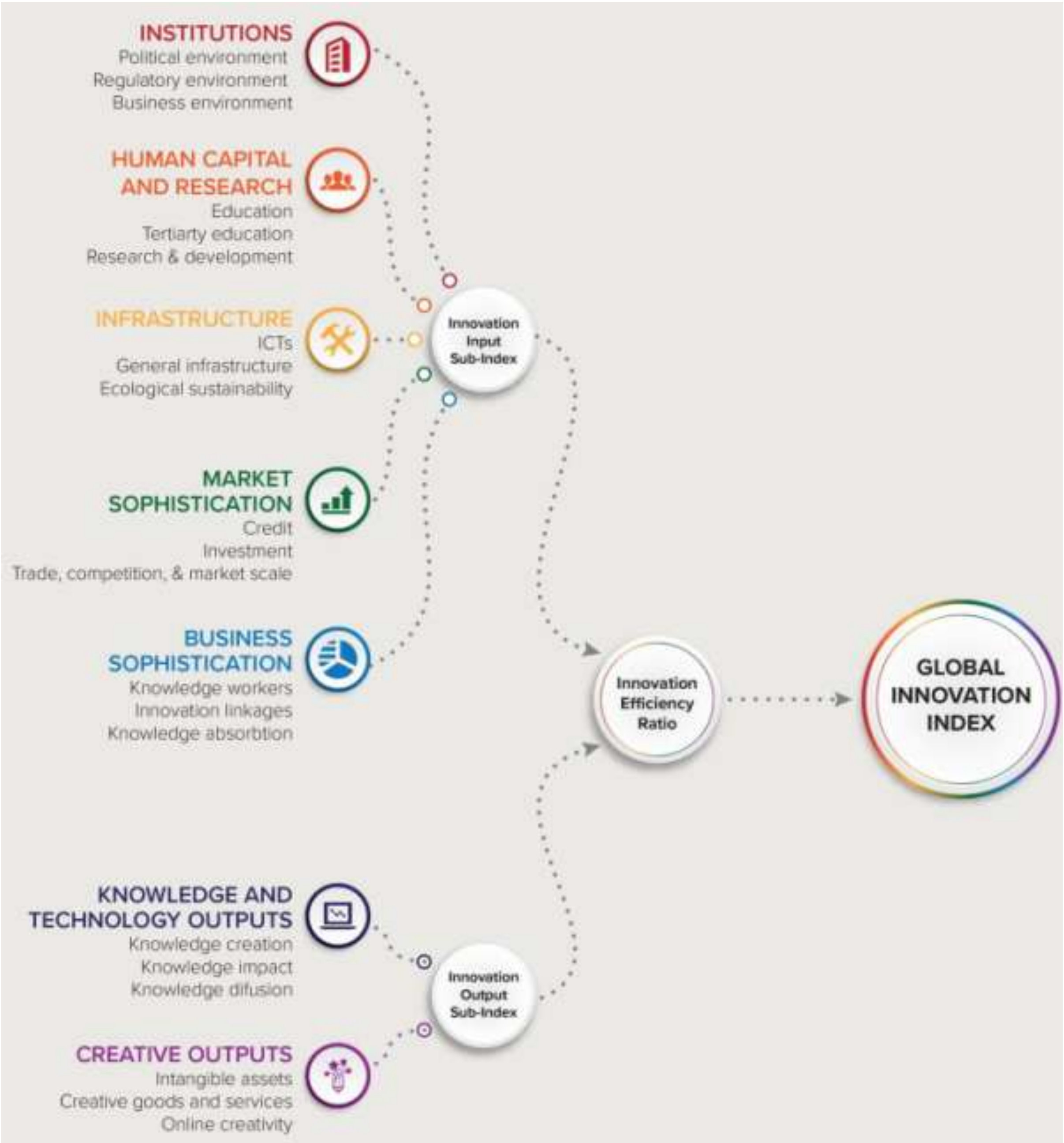

Figure 1: Global Innovation Index Framework (Global Innovation Index, 2018)

As seen on the graphic above, the indicators from Global Innovation Index framework are divided into two sub-indices, according to a segmentation well-established in the literature, creating two groups of indicators: innovation inputs & outputs (Archibugi et al., 2009). Each pillar is divided into three sub-pillar consisting of various indicators. Innovation input sub-index consists of 5 input pillars aimed to capture elements of the national economy that enable innovative activities. Innovation output sub-index contains results of innovative activities within each country. It is noted in the report that even though the output sub-index includes only 2 pillars, it has the same weight in calculating the overall GII scores as the input sub-index. The innovation efficiency ratio is calculated as the ratio of the innovation output sub-index over the input sub-index to show how much innovation outputs a given country is getting for its inputs. Lastly, the overall GII score is counted as a simple average of the input and output sub-indices (Global Innovation Index, 2018). Below, there is a description, taken from the report, of the whole framework and indicators is consists of.

The first sub-index has five pillars, which define aspects of the environment favorable for creation and diffusion of innovations: Institutions, Human capital & research, Infrastructure, Market sophistication, Business sophistication. The institutions pillar's purpose is to capture the institutional framework of a country. It includes variables related to political, regulatory, and business environment. The political environment consists of two indices: one that reflects the likelihood that a government might be destabilized, and second one that reflects the quality of public and civil services, policy formulation, and implementation. The regulatory environment is based on 2 indices that aim to capture the ability of the government to formulate and implement cohesive policies, which promote the development of the private sector, and evaluate to which extent the rule of law prevails (in aspects such as contract enforcement, property rights, the police, and the courts). The 3rd indicator evaluates the cost (in salary weeks) of terminating a redundant worker. The business environment sub-pillar examines 2 aspects that directly affect private entrepreneurial endeavors by using the World Bank indices on the ease of starting a business, and ease of resolving insolvency. This year the framework has changed and the indicator measuring ease of paying taxes was dropped.

The second pillar of Global Innovation Index is called human capital & research. It tries to capture the human capital of countries, as the level and standard of education and research activity in a country are said to be prime determinants of innovation capacity. The first sub-pillar includes a set of indicators that capture achievements at the elementary and secondary education through education expenditure, school life expectancy, and government funding per pupil. The quality of education is measured through pupil/teacher ratio, as well as, the results of the OECD Programme for International Student Assessment (PISA), which examines 15-years old students in reading, mathematics, and science. Higher education is said to be crucial for economies to move up the value chain beyond simple production processes and products. Thus, the sub-pillar on tertiary education captures tertiary enrolment, with priority on sectors traditionally associated with innovation: percentage of tertiary graduates in science, engineering, manufacturing, and construction. Other indicators include inbound mobility of tertiary students, which has a crucial role in the exchange of ideas and skills necessary for innovation. The last sub-pillar under the human capital and research concerns R&D and aims at measuring the level, and quality of R&D activities. Its indicators focus on researchers (full-time equivalent), gross expenditure on R&D, R&D expenditure of top 3 firms, and the quality of scientific & research institutions, measured by the average score of the country's top 3 universities in the QS World University Ranking.

It is said that good and ecological communication, transport, and energy infrastructure facilitate the creation and exchange of ideas, services, and goods, contributing to the innovation system through increased productivity and efficiency, better access to markets, lower transaction costs, and sustainable growth. Thus, the third pillar regarding infrastructure includes sub-pillars on Information and communication technologies (ICTs), General infrastructure, and Ecological sustainability.

The ICTs sub-pillar includes 4 indices on ICT access, ICT use, online service by governments, and online participation of citizens. The sub-pillar on general infrastructure involves the average electricity output in kWh per capita, a composite indicator on logistics performance, and gross capital formation. The sub-pillar on ecological sustainability includes 3 indicators: GDP per unit of energy use, Environmental Performance Index of Yale & Columbia Universities, and the number of issued certificates of conformity with standard ISO 14001 on environmental management systems.

The next pillar is called market sophistication and consists of 3 sub-pillars structured around market conditions and total level of transactions. In broad terms, the pillar is based on the assumption that well-functioning markets contribute to the innovation environment through competitive pressure, efficiency gains, economies of transaction, and by allowing supply to meet demand. Markets that are open to foreign trade and investments have the additional effect of exposing domestic firms to best practices around the globe, which is critical to innovation through knowledge absorption and diffusion.

The credit sub-pillar includes a measure on the ease of getting credit, capturing to which degree collateral and bankruptcy laws facilitate lending by protecting the rights of borrowers and lenders, as well as, the rules and practices that affect the coverage, scope, and accessibility of credit information. Transactions in this sub-pillar are given by the total value of domestic credit and, in attempt to make the model more applicable to emerging markets, by the gross loan portfolio of microfinance institutions. The availability of credit and the environment that supports investment, access to the international market, competition, and market scale are said to be critical for businesses to prosper, and for innovation to occur.

The investment sub-pillar involves the ease of protecting minority investors index and 2 indicators on the level of transactions, one analyzes whether market size is matched by market dynamism, and second one is a hard data metric on venture capital deals. The last sub-pillar under market sophistication tackles trade, competition, & market scale. Its first indicator measures the average tariff rate weighted by import shares and shows the market conditions for trade. As there is no hard data collected on competition, the second indicator is a survey question that reflects the intensity of competition in local market. The last sub-pillar takes into consideration the impact that the size of an economy has on its capacity to introduce and test innovations on the market by domestic market scale, incorporated in 2016, measured by economy's GDP.

The last pillar regarding innovation inputs is named business sophistication. It assesses how conducive firms are to innovation activity. The human capital & research pillar was based on the premise that the accumulation of human capital through education, particularly higher education, and the prioritization of R&D activities, is a fundamental condition for innovation to take place. This logic is taken one step further in the business sophistication pillar, with the assertion that businesses foster their productivity, competitiveness, and innovation potential with the employment of highly qualified professionals, and technicians.

The first sub-pillar includes 4 quantitative indicators on knowledge workers: employment in knowledge-intensive services, the availability of formal training at the firm level, R&D performed by business enterprise (GERD) as a percentage of GDP, and the percentage of total gross expenditure of R&D financed by business enterprises. Additionally, this sub-pillar includes an indicator related to the percentage of females employed with advanced degrees, which provides a glimpse into the gender labour distributions of nations, and offers more insights to the degree of sophistication of the currently employed local human capital.

Innovation linkages and public/private/academic partnerships are considered essential to innovation. In emerging markets, industrial or technological clusters and networks have become more wealthy, in sharp contrast to the poverty that may prevail in the rest of the territory. The innovation linkages sub-pillar relies on both quantitative and qualitative data related to business/university collaboration on R&D, the prevalence of well-developed clusters, the level of gross R&D expenditure financed by abroad, and the number of deals on joint ventures & strategic alliances. Additionally, the total number of Patent Cooperation Treaty (PCT), and national office published patent family applications filled by residents in at least 2 offices are proxies for international linkages.

The rationale behind sub-pillars on knowledge absorption (enabler) and knowledge diffusion (result), two sub-pillars designed to be mirror images of each other, is constructed in a way that together they reveal how good each economy is at absorbing and diffusing knowledge. Sub-pillar on knowledge absorption includes 5 metrics linked to sectors that are key to innovation: intellectual property payments as a percentage of total trade, high-tech net imports as a percentage of total imports, imports of communication, computer and information services as a percentage of total trade, and the 3-years average of net inflows of foreign direct investment (FDI) as a percentage of GDP. In 2016, the percentage of research talent in business was added, in order to strengthen the sub-pillar and provide a measure of professionals engaged in the conception or creation of new knowledge, methods, products, processes, and systems, including business management.

Innovation outputs are the results of innovative activities within the economy. This sub-index is built upon 2 pillars: knowledge & technology outputs and creative outputs. The knowledge & technology outputs pillar covers all variables that are traditionally thought to be the direct results of inventions and/or innovations. These variables refer to the creation of knowledge, its impact, and diffusion. It includes 5 indicators, representing the results of inventive and innovative activities: patent applications filed by residents both at the national patent office and internationally through PCT, utility model applications filed by residents at the national office, scientific & technical articles published in peer-reviewed journals, and economy's number of articles (H) that have received at least H citations.

The second sub-pillar, regarding knowledge impact includes statistics that represent impact of innovation activities at the micro- and macroeconomic level: increases in labour productivity, density of entry of new firms, number of certificates of conformity with standard ISO 9001 on quality management systems issued, spending on computer software, and the measure of high- & medium-high-tech industrial output over total manufactures output.

The third sub-pillar, regarding knowledge diffusion, is the mirror image of the knowledge absorption sub-pillar. It includes 4 statistics all linked to sectors key to innovation: intellectual property receipts as a percentage of total trade, high-tech net exports as a percentage of total exports, exports of ICT services as a percentage of total trade, and 3-year average of net outflows of FDI as a percentage of GDP.

The last pillar of the Global Innovation Index framework is related to creative outputs. It is based on a statement that the role of creativity for innovation is largely underappreciated in innovation measurement and policy debates, therefore the GII has emphasized measuring creativity as part of innovation output sub-index. Its first sub-pillar considers intangible assets and includes statistics on trademark applications by residents at the national office, industrial designs included in applications at a regional or national office, and two survey questions regarding the use of ICTs in business and organizational models, which are new areas that are increasingly linked to process innovations in the literature.

The 2nd sub-pillar of creative outputs is based on creative goods and services. It has been improved throughout recent years. In 2014, global entertainment and media output was added in an attempt to include broader sectoral coverage. In 2017, the indicator on exports of audio-visual and related services was renamed "Cultural & creative services exports", and expanded to include information services, advertising, market research, public opinion polling, as well as, other personal, cultural, and recreational services, as a percent of total trade. The 3rd sub-pillar measures national feature films produced in a given country (per capita), printing and recorded media output (as a percentage of total manufactures output), which has also been changed to precisely capture printing & media outputs, and exclude paper industry outputs. This sub-pillar also measures creative goods exports (as a percentage of total trade), which is aimed to provide the overall sense of the international reach of creative activities in the country.

The 3rd sub-pillar on online creativity includes: generic and country-code top-level domains, average yearly edits to Wikipedia, scaled by population aged 15-69 years old, and mobile app creation, which is scaled by GDP (PPP). The indicator on mobile app creation replaced the indicator on video uploads on YouTube in the last edition of GII. Mobile apps are considered to represent the global commerce in digital goods, and therefore provide insight into how innovation, production, and trade of digitalized creative products, and services evolve in an innovation-based economy.

It is worth noting that the dataset does not include the full numbers that represent the indicators, just the ranks for each country in given criteria. So, most variables include numbers from 1-126, less if there is missing data for some economies.

5. Theory

This section will introduce various theories applied in this thesis. Firstly, the concept of innovation, together with different definitions for it, will be introduced, as it is a central theme of this project, and of the analyzed Global Innovation Index. Subsequently, the definition of innovation that Global Innovation Index is based upon will be presented, together with interesting changes that have been made overtime. After defining innovation and comparing differences and updates of description of the concept, the bases for innovation measurement will be presented, followed by distinction of various indicators and variables used to measure innovation, and challenges in their measurement. Furthermore, the systems approach to analyzing innovation will be shown in order to gain a fundamental understanding of reasoning behind measuring innovation on a macroeconomic level. The subchapter will begin with introducing systems of innovation on different aggregation level but the main focus will be placed on national systems of innovation, as the analysis on a country level is of the main interest of this project. The final theoretical considerations will be based on the concept of competitiveness, especially between countries, to get an understanding of how the concept is defined and how countries gain a competitive advantage.

5.1 Innovation

The term innovation has gained a lot of interest among researchers in various areas throughout recent years. A single search for the keyword "innovation" in Google Scholar returns 4,070,000 results. It is also a key concept of Global Innovation Index and a major focus of MIKE study programme.

The *"Innovation - a New Guide"* working paper includes a graph presenting the number of publications with innovation in the title from 1956 until 2008, with the data taken from the British Library and the ISI Web of Science. As it seems to be a good measure of how popular the field of innovation has become in the research community, I decided to explore what was the most important time period, in the last 30 years, when the term "innovation" in scientific literature was in its peak.

The data for this bibliographic analysis was obtained from Scopus, which is one of the largest bibliographic databases with abstracts and citations from research papers in different fields of science, where the search was limited to only include exact keyword "innovation", and to return only publications in English. Documents were sorted in descending order by the number of citations and as a result of limit for downloading a single file, only the first 2000 records were included in the analysis.

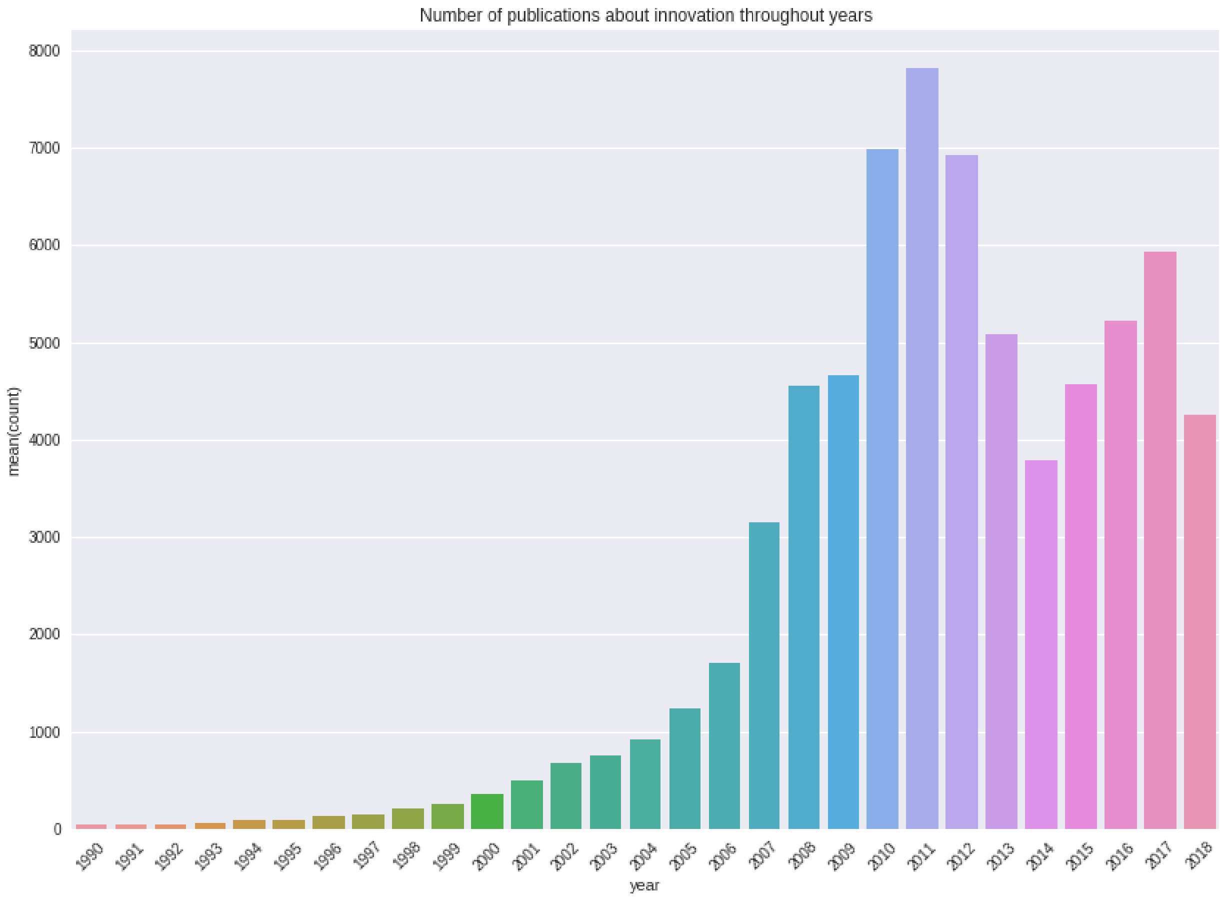

Figure 2: Number of publications about innovation since 1990

Fagerberg (2013) presents *Innovation Studies: The Core Literature* to highlight what are the ten most important contributions to the core literature on innovation studies and the newest title included in the ranking is from 1993. Together with the graph above, this may lead to a conclusion that even though the most important concepts were defined in the 20th century, the popularity of innovation studies actually started increasing in the last 20 years. It is clear from the graph that this area of research was gaining interest quite steadily until 2006 and then started growing more rapidly to reach its peak in 2011, with the exact number of publications equal to 7812. In the years 2012-2014 the number of publications related to innovation studies was decreasing and then started growing again.

Innovation studies are defined as the analysis of how innovations emerge and diffuse, which factors are influencing these processes, and what are their social and economic consequences. The origins of this field of research date back to the post Second World War period, when researchers started addressing questions concerning the role played by innovation that influences economic and social change. It is said that even though innovation is a very important topic, it has not always received the scholarly attention that it deserved. There were only a few publications about the topic until the early 1960s. One of the main exceptions, in the early days of social science, was the work of Joseph Schumpeter, who was trying to combine insights from economics, sociology, and history into an original approach to the study of economic and social change in the long run, particularly focusing on the crucial role played by innovation and the factors influencing it (Fagerberg, 2013).

In *"Theory of Economic Development"* innovation is seen as the major mechanism behind economic dynamics. The generator in the system is the individual entrepreneur, which introduces innovations in markets and creates new enterprises. After the pioneers, imitators follow and the profits created by the original wave of innovation arise gradually. In *"Capitalism, Socialism and Democracy"* Schumpeter stated that the main source of innovation is not an individual entrepreneur but big firms with many experts working together in R&D

departments, searching for new technological solutions. This distinction between the two ways of presenting the motor of innovation has led scholars to refer as Schumpeter Mark I and Schumpeter Mark II (Lundvall, 2017).

Joseph Schumpeter has also defined an important distinction between innovation and invention. Invention is described as the first occurrence of an idea for a new product or process and subsequently, innovation is the first attempt to implement it in practice. In order to turn an invention into innovation, several different types of capabilities, resources, skills, and knowledge need to be combined. There is also a requirement to possess necessary skills, production knowledge, market knowledge, facilities, sufficient financial resources, and a well-functioning distribution system. Therefore, the role of an innovator, such as a person or organizational unit responsible for combining these required factors, that was called an entrepreneur by Schumpeter, may be different from the role of an inventor (Fagerberg, 2013).

In some cases, inventions and innovations are also being so closely linked together that it becomes difficult to differentiate one from another. There are also instances, where commercialization of one idea is impossible until another invention, new knowledge, or technology arises, that would be complementary for the original idea and necessary for its successful implementation. As a result, there are cases of new products, which are seen as a single innovation, but in fact are a result of a lengthy process that involves many interrelated innovations (Fagerberg, 2013).

Another important distinction is made between innovation and imitation. Imitation is the replication of innovation and it also has a great social and economic significance, as without imitation the impact of innovation would be much less significant, if any at all. However, even though the distinction between innovation and imitation is transparent enough in theory, it is often difficult to apply it in practice. One of the reasons for this difficulty is that some products or processes might be well known in one context but completely new in another context. The main discrepancy is when the same innovation is introduced by different actors in diverse contexts. It is not clear whether both actors should be characterized as innovators, because the circumstances or markets are different, or only the first one should be an innovator and the second one should be named imitator. According to Schumpeter's distinctions, the right way of differentiating between the two should follow the second example. On the other hand, the European Union's Community Innovation Survey (CIS) adopts the other definition of innovation versus imitation, while the same concept introduced in a different context would count as an innovation (Fagerberg, 2013).

Innovations are commonly classified according to their types. Based on the Schumpeter's classification, 5 different types were distinguished: introduction of new products, new methods of production, development of new sources of supply, the exploitation of new markets, and new ways to organize businesses. In economics, the most focus and attention have been placed on the first two types. The term product innovation has been used to characterize the occurrence of new or improved goods and services, while the definition of process innovation involves improvements in the ways to produce these good and services (Fagerberg, 2013).

Another way of classifying innovations, also based on Schumpeter's work, is a classification depending on how radical the innovations are. From this perspective, three different types were distinguished. Continuous improvements are characterized as incremental or marginal innovations and more breakthrough innovations, such as the introduction of a totally new type of machinery in a specific industry, are called radical. Radical innovations can be also defined as innovations which transform existing markets or industries, and based on which many incremental innovations are developed. The last type of innovation is described as technological revolutions, which consist of a cluster of innovations that together can have a very far-reaching impact in a whole range of industries or the economy as a whole. Recently, they have been also called general purpose technologies. Schumpeter's work was particularly focused on radical innovation and technological revolutions, which he believed were of greater importance (Fagerberg, 2013)(Lundvall, 2017).

Schumpeter has also argued that economic development is driven by innovation, which is a dynamic process where new technologies replace the previous ones. This process was named creative destruction and has been characterized as revolutionizing the structure of a market, especially from within, by creating innovations while destroying the old technologies. Based on Schumpeter’s work, Clayton Christensen developed this theory further, suggesting the term disruptive innovation for innovations that undermine the position of existing business models, through the continuous exploitation of new markets or market niches (Oslo Manual, 2005)(Fagerberg, 2013). Another distinction of innovations is the concept of open versus closed innovation, largely popularized by Henry Chesbrough.

5.2 Innovation according to Oslo Manual

As it can be seen from subchapter above, there are many different types of innovations and approaches to studying or defining the concept. Firstly, in order to conduct analysis of the innovative performance of countries, it is required to establish a common understanding of innovation. It is said that in order to measure innovation and its subsequent economic outcomes, precise definitions of innovation and innovation activities are needed beforehand (Oslo Manual, 2018). The Global Innovation Index adopts a definition of innovation, originally elaborated in the Oslo Manual 2005, developed by the European Communities and the Organisation for Economic Co-operation and Development (OECD):

“An innovation is the implementation of a new or significantly improved product (good or service), a new process, a new marketing method, or a new organizational method in business practices, workplace organization, or external relations“
(Global Innovation Index, 2018)

The first edition of the Oslo Manual was published in 1992, as a synthesis based on a broad group of innovation surveys from the late 1980s, to provide a standardized framework for collecting data about technological product and process innovation in manufacturing industries. The framework was updated in 1997 to include innovation in services. The third edition of the Oslo Manual was published in 2005, after undergoing a number of substantial changes, aiming to keep innovation measurement alongside policy needs and including developments in the innovation literature. The most important changes contained a broadened definition of innovation, which included marketing and organizational innovations; an expanded coverage of knowledge flows as well as the role of linkages in the innovation process; and an adaptation to reflect increased awareness of innovation in less R&D intensive industries, including services and low-tech manufacturing (Bloch, 2007).

Broadening the innovation concept is said to open up for examination of a wider range of innovation activities, which allows the analysis of non-technological types of innovation and the examination of links among different types of innovation activity. It is argued in a number of areas of innovation literature, including innovation systems approach, that there is a need to broaden the innovation concept. Therefore, the introduction of marketing and organizational innovations in the 3rd edition of Oslo Manual was an important step in this direction (Bloch, 2007).

The minimum requirement to call some product, process, marketing or organizational method an innovation, according to the definition adapted in the Global Innovation Index, is that they must be new or significantly improved. This includes products, processes, and methods, which are both newly developed, as well as, the ones that have been adopted from other firms or organizations (Oslo Manual, 2005).

Oslo Manual introduces interpretations for each type of four innovations described in the main definition. Namely, product innovation is described as introducing a good or a service, which is new or significantly improved taking into consideration its characteristics or intended uses. These changes include significant improvements in components and materials, technical specifications, incorporated software, user-friendliness, or in other functional characteristics. Similarly, process innovation is defined as the implementation of a new or significantly improved production or delivery method, which includes significant changes in techniques, equipment, or software (Oslo Manual, 2005).

Organizational innovation is interpreted as the implementation of a new organizational method in the firm's business practices, organization of a workplace, or the external relations. The theoretical bases for this innovation are formed on the literature on organizational innovation (for instance the chapter by Alice Lam in The Oxford Handbook of Innovation), which is centered around the role of organizational structures, learning processes, and the adaptation to changes in both technology and environment. It is said that a firm's organizational structure can affect the efficiency of innovation activities and some organizational structures are better suited to particular environments. For instance, in order to generate more radical innovations, a looser and more flexible form of organization can be more effective, as it allows employees to have greater autonomy to make decisions and define their responsibilities. On the other hand, a greater degree of organizational integration might improve the planning, coordination, and implementation of innovation strategies, which works particularly well in industries characterized by incremental changes in knowledge and technologies (Oslo Manual, 2005). Therefore, it is important to take organizational changes into account while analyzing different types of innovations on a microeconomic level.

The last type of innovation described in Oslo Manual (2005) is a marketing innovation. It is defined as the implementation of a new marketing method, which involves significant changes in product design or packaging, placement, promotion, or in pricing. Many firms allocate substantial resources to product design and packaging, where their main activities can involve market research, testing, and the use of new knowledge. The same principles are also true for marketing practices, through which many firms focus on marketing and design of their products or services in order to improve their competitive performance. The design of products and marketing practices may both be aimed at improving the presentation and promotion of offered goods. Drawing on these premises, the definition of marketing innovation in Oslo Manual is mostly based on normative marketing theories that are focused on the implementation of marketing practices. In this case, it is the Marketing Mix Model, introducing 4P's model for marketing strategies: product, price, promotion, and placement (Bloch, 2007)(Oslo Manual, 2005).

It is said that innovation activities can lead to both development and implementation of innovations in the short term, and to further improvements in the innovative capacity of the firm. Firms learn by developing and implementing innovations, gain valuable inputs from interactions and marketing activities, and improve their innovative capacities through organizational changes (Oslo Manual, 2005).

The definition of innovation, used within the Global Innovation Index framework, was taken from the 3rd edition of Oslo Manual (2005) but in 2018 it was revised again the fourth edition of guidelines for collecting, reporting, interpreting, and using data on innovation was published. The most updated version of Oslo Manual has a new definition of innovation, which might be incorporated to Global Innovation Index in the future. The new interpretation is as follows:

"An innovation is a new or improved product or process (or combination thereof) that differs significantly from the unit's previous products or processes and that has been made available to potential users (product) or brought into use by the unit (process)"

(Oslo Manual, 2018)

The main difference between the two versions of definition of innovation is that the interpretation from 2005 version included new organizational methods, marketing, goods, and services. It is stated that this major change of reducing 4 types of innovations (product, process, organizational, marketing) to 2 main types product innovations and business process innovations was based on cognitive testing work (Oslo Manual, 2018).

Moreover, the updated description highlights that the product or process should differ significantly from previous ones, which seems like incremental innovations are excluded. In the further clarification on types of innovations according to their novelty and impacts, the manual explains that the basic requirement to call a new product or process innovation is that it must be significantly different from previous products or business processes. But "significantly different" is a subjective term and will vary between firms. Thus, its understanding will be different, based on interpretation, various norms, or firm's capabilities, and contexts. There are also different measures of novelty, innovativeness, and economic impacts. Those measures involve an indication whether an innovation is new to the firm, new to the market or new to the world. Those indicators can also include firm's expectations of the potential to transform the market in which it operates or the firm's expectations of the potential to improve its competitiveness (Oslo Manual, 2018).

In order for a new idea, method, model or prototype to be considered an innovation, it needs to be implemented, which is also a defining and distinguishing characteristic of innovation. The implementation is included in both definitions from different versions of Oslo Manuals. Implementation requires that organizations will make systematic efforts to ensure that innovation is accessible to potential users. Defined access can be both to the organization's own processes and procedures or for external users to products. As a minimum requirement, innovations must contain characteristics that were not previously made available by the relevant organization to its users. These features may, but do not have to be, new to the economy, society, or a particular market. It is possible for an innovation to be based on products and processes that were already used in other contexts, for example in other geographical or product markets. In this case, innovation is an example of diffusion. Diffusion of innovation can generate substantial economic and social value and, as a consequence, is important for policy-makers. Additionally, implementation is not the final step of innovative activities. The activities that follow-up, for example reviewing innovations after their implementation, can also result in minor improvements or radically new innovations, through a fundamental redesign of products, processes or methods. Some of those follow-on activities could potentially result in innovations on their own. Moreover, post-implementation reviews can lead to the abandonment of innovations in some cases (Oslo Manual, 2018).

The revised Oslo Manual also includes the definition of innovation activities, which are characterized as activities involving all developmental, financial, and commercial actions undertaken by a firm that are intended to result in an innovation (Oslo Manual, 2018).

It is interesting to look back into the definition of innovation from the second version of Oslo Manual. The description from 1997 also includes only process and products innovations, although they are called "technological product innovation" and "technological process innovation". Subsequent innovation surveys have shifted their focus from the use of the word "technological", to make the definitions more applicable to innovation in services and less manufacturing oriented. Therefore, consecutive definitions of innovations do not include the word "technological" (Bloch, 2007).

Despite of this main difference in nomenclature, the definition of innovation from 1997 is closer to the newest revision of Oslo Manual, than the version from 2005. The most interesting is that the revision from 2005 was based on a statement that innovation activities of firms are clearly more than product and process innovation and the definition should be modular, so that the analysts can still focus on the narrower concept of product and process innovation (Bloch, 2007). Regardless that this was a premise for the update of interpretation of this term, the newest edition of Oslo Manual came back closer to 2nd edition, after conducting empirical tests concerning the respondents to surveys based on the broadened definition that included four forms of innovation.

Even though operationalizing the broader definition of innovation was supposed to make the analysis easier, it created a number of additional challenges. One of them was creating individual and distinguishable definitions for each type of innovation, included in the definition, and adapting other topics of surveys to the broadened innovation concept (Bloch, 2007).

The first issue arose already in the phase of forming a definition of organizational innovation. It is related to defining when organizational changes should be considered innovations. First criterion is the objectives of the change. Organizational innovations share the same objectives as process innovations in the way that they may aim to enhance production or delivery capabilities through improvements in efficiency, productivity, or quality, or by facilitating the manufacturing of new products. However, following the description by Lam, a broader view of the innovative firm also includes capabilities to create new ideas, to use existing knowledge in new ways, and to transform knowledge into marketable products. Based on that it can be concluded that organizational innovations should not only be linked to changes which aim to improve the production and delivery of existing goods and services, but also to the ones which aim to improve capabilities for developing new products and processes. Organizational innovations are defined as consisting of minimum 1 of the 3 elements: business practices, workplace organization, and external relations. Business practices are defined as routines or procedures for the conduct of work, they can range from practices for sharing knowledge to the sets of procedures involved in management systems. The second element, workplace organization, involves organizational structures and the distribution of responsibilities, and decision-making. External relations involve the organization of relationships with other firms or public research institutions (Bloch, 2007).

Another important problem is how to identify significant changes in characteristics in order to apply similar criteria to organizational and marketing innovations as for product and process innovations. There is a requirement that marketing and organizational innovations should consist of methods that have not been used before within the firm, and are a result of strategic decisions taken by management. But a significant challenge arises in defining the 4 types of innovation when it comes to distinguishing among them in borderline cases, as sometimes it is not considered feasible or desirable to make clear-cut distinctions among different types of innovations. Distinctions often depend on the specific characteristics of a firm's innovations, and on the nature of the firm's business in general. For instance, it is often argued that production and consumption are frequently blurred in services, which is making it difficult to distinguish between product and process innovations. Service firms might be able to identify what type of innovation is based on the characteristics of the actual changes made. If the characteristics of the service that the customer experiences are new or significantly improved, it would be classified as product innovation. On the contrary, if the firm makes significant improvements to how the service is provided, possibly without changing the service's characteristics, it would classify as process innovation. However, in many cases it is still difficult to make a distinction among the firm's activities, including product development, production and delivery processes, organizational practices, and marketing (Bloch, 2007).

Additionally, issues related directly to organizational and marketing innovations started to arise. It is said that analysis of commercialization is not limited to innovation in marketing methods, but it also includes the variety of marketing and organizational practices, which firms use for example to conduct market research and to utilize market knowledge in development of their products. One potential problem is that there might be confusion between current practices and innovations. Therefore, even though a firm could state that there have been changes in existing organizational practices, there is no explicit criteria for these changes to be considered as innovations, in contrast to the definitions of organizational and marketing innovations (Bloch, 2007). As a result, there might disorientation to some degree, caused by the inability to distinguish what should be classified as organizational and marketing innovations. It was also a part of explanation why the definition of innovation was simplified again in the newest edition of Oslo Manual and reverted to only two types of innovations. Empirical research has proven that managers find it difficult to differentiate between organizational and process innovations. Thus, organizational innovations in the updated version of manual are incorporated under one type of business process (administration and management), which includes activities such as strategic management and human resource management (Oslo Manual, 2018).

5.3 Measuring innovation

In order to analyze how the innovation rankings are constructed, it is important to understand how to measure innovation, what are the most common indicators addressing this problem, and what are the biggest challenges in measuring innovation. Some issues related to defining and distinguishing various types of innovations were already mentioned in previous subchapter but ways of collecting data, and constructing indicators also have big impact on analysis, therefore they will be introduced further in this section.

The key assumption of the revised Oslo Manual is that innovation can and should be measured. This requires both an understanding of what needs to be measured and the awareness of what could be reliably measured. The Oslo Manual is a publication which contains the definition of innovation that the Global Innovation Index is based upon. But it is also a guideline for collecting and interpreting data on innovation, as well as, a tool to support national statistical offices and other producers of innovation data in designing, collecting, and publishing various measures of innovation. It is said that a robust measurement of innovation, together with the use of innovation data in research, can both help policymakers to understand better the economic and social changes, assess whether the contribution of innovation to social and economic goals is positive or negative, monitor and evaluate the effectiveness and efficiency of different policies (Oslo Manual, 2018).

The main conceptual foundations for innovation measurement are primarily derived from economics and management. The management perspective on innovation aims to explain how innovation can change firm's position on the market and how to generate innovative ideas. Economic perspective, on the other hand, examines why organizations innovate, the forces which drive innovation, factors that hinder it, and the macroeconomic effects of innovation on the whole market, industry, or economy (Oslo Manual, 2018).

It is said that the needs of users drive the construction of a system to measure and report innovation, as well as, the subsequent production of innovation statistics, data, indicators, and in-depth analyses of innovation activities. There is a common and widespread interest in understanding what drives firms, communities or individuals to innovate, and what are the factors that influence their innovation activities the most. The relevance of innovation data for understanding innovation processes and drivers varies between countries, industries, and different institutional settings. The usefulness of data on innovation also depends on the ability to connect it with other types of data. There are 3 main users of innovation data: managers, academics, and policymakers. The most important purposes of collecting data about innovation are: managers can use them as an input for forming firms' strategies, academics or researchers can conduct theoretical analysis, and it is a source of information for forming public policies (Archibugi et al., 2009). The data needs for all types of users are similar, with the main interests in obtaining comparable data across industries, regions, and time; providing data on the factors that enable or hinder innovation; linking innovation data to other relevant data, such as administrative registers or data on individual users of innovations; keeping up with changes in the nature of innovation, such as open innovation; enabling analyses of innovation impacts on different levels of aggregation, for example on innovative organisations, regions, or nations (Oslo Manual, 2018).

The choice which methods should be adopted to measure innovation depends on the quality of the collected data and its intended use. A strategy for measuring innovation must address several issues, such as the choice of an approach, the choice whether to collect qualitative or quantitative data, sources of data, and the responsibility for data collection. The structure of a measurement strategy can change over time, as the user needs and types of data, which can be collected, evolve as a result of new opportunities or challenges. Additionally, various measurement approaches can complement each other. The value of innovation measurement can often be improved by combining several approaches and by creating opportunities for data linkages, and follow-on analysis (Oslo Manual, 2018).

A choice that has to be made before conducting analysis is whether to collect qualitative or quantitative data. For instance, academics and policymakers prefer quantitative data for most research purposes. However, respondents of surveys find it difficult and demanding to report quantitative, interval data for innovation activities or outcomes, such as innovation expenditures, income generated by innovations, the number and length of collaborations, the number of IP registrations or applications, personnel, etc. Additionally, many concepts about innovation are difficult to quantify, partially because the records and management systems of firms do not align with innovation concepts, or just because the concepts only apply to specific contexts. Those qualitative measures for innovation activities that cannot be collected on an interval level can be collected and codified by using nominal or ordinal data. This could include the importance of different information sources or categories characterized by the frequency with which these sources are accessed. This type of qualitative data is possible to use in econometric analyses and to construct indicators. There is considerable scope for using unstructured qualitative data to formulate statistics. Examples include self-reported descriptions of firm's most important innovation, descriptions of innovation strategies in company or organizational reports. Researchers used to codify this type of data manually, although currently there is a shift towards using machine-based algorithms that use natural language processing techniques to process textual data (Oslo Manual, 2018).

It is said that not all indicators that are useful for analysis of a single country are suitable for benchmarking across different economies, mainly due to linguistic, cultural, and contextual differences. An innovation measurement framework covers a predefined scope, such as a sector of interest, jurisdiction or geographic area, where data has been collected, set of relevant phenomena of interest necessary for understanding innovation, and measurement strategies (Oslo Manual, 2018).

In order to determine if a dataset and indicators are well-suited for policymakers, the goals of public policy need to be identified beforehand, to ensure that the measurement framework matches the policy needs. Policymakers not only influence the types of data that are collected, but they can also have an impact on the extent and quality of obtained data through supporting or funding new data collection, or linking data to existing sources. The base of users for innovation statistics is evolving over time, as the statistical data on innovation is proven to be more or less relevant for informing decisions in specific circumstances, or as a new data become available. Innovation data is relevant to a wide array of policy areas, which includes general macroeconomic management, public services, taxation, and environmental policies. The data on innovativeness can be particularly informative for the study of structural policies because of the high degree of persistence of many innovation-related behaviors. This also means that some types of innovation data do not need to be collected frequently, although the timely data would be valuable in the presence of rapid structural change or at times of economic or financial crises (Oslo Manual, 2018).

Based on various factors present in the measurement framework, it is possible to identify what are the characteristics which distinguish each country in relation to analyzed phenomena, as well as, how does it compare to partners and competitors. When used in the right way, analysis of aggregate indicators can help politicians to take important decisions concerning international cooperation agreements, in other words, to choose partners according to their complementary competences. Moreover, since a group of countries establishes common objectives, the statistical tools can help to verify in which measures each member state is contributing to achieving the objectives. International comparisons of different economies also allow researchers to identify convergences or divergences across countries. The analysis of convergence is particularly important and interesting for the European Union. In a moment when various member states intend to strengthen their cohesion and to adopt a common strategy for innovation, it becomes relevant for the authorities and policymakers to identify the contribution provided by each member state (Archibugi et al., 2009).

It is common practice by many researchers interested in measuring innovation to firstly set a level of aggregation, for example on a firm, industry, region, or country level, and compare the results between similar entities. This kind of benchmarking is believed to be useful for many purposes. However, innovations are inherently difficult to measure precisely (Fagerberg, 2013).

The rankings, constructed based on innovation data, provide preliminary information for policy actions. However, it is important to note that the various components of aggregated indicators are treated as substitutes in terms of statistical importance, while the assumption does not have anything to do with reality, especially taking into consideration policy perspective. If a country is lacking in one area of development, for instance infrastructure, it is possible to improve its rank by compensating it with another area, for example increase in academic production measured by bibliometric indicators, but it will not improve the situation within the country. Thus, the aim of policymakers should not be to increase the value of indicators but rather to improve the economic and social conditions those indicators are expected to capture. Statistical data collected for those rankings can be useful to identify the main obstacles to national development, leading to addressing appropriate public policies. In order to do so, it is needed to distinguish the indicators from related phenomena, as for example numbers of publications and patents are means, not ends, and do not necessarily result in knowledge creation or diffusion (Archibugi et al., 2009).

In the past, in order to measure innovation, economists and policymakers used to put emphasis on R&D-based technological product innovation, which was principally produced in-house, particularly in manufacturing industries. The process leading to developing such innovation was conceptualized as closed, internal, and localized. There was a need for radical innovation and technological breakthroughs in order to stay competitive, which implied the existence of leading and lagging countries, while the low or middle-income economies were only catching up (Global Innovation Index, 2018). Traditionally, investing in R&D has been regarded as one of the main strategies to secure technological potential and, as a result, also innovations and economic growth. R&D investment was said to increase the possibility of achieving a higher standard of technology in firms or regions, which allowed introducing new and superior products or processes, resulting in higher levels of reported income and growth. It has also been proven that there is a relationship between investment in technology and R&D expenditure, which results in increases in productivity and growth. Therefore, the relationships between these variables (from R&D investment to technological potential and from technological potential to innovation and growth) is considered to show the path, which policymakers should follow, in order to secure economic growth in any given region. Over the last few years, public administrations have followed this framework and, thus, policies aimed at increasing the role and importance of technology, by promoting research activity through the increase of public investments in R&D, have been frequently formed in their respective territories (Bilbao-Osorio & Rodríguez-Pose, 2004).

One of the most frequently used measures for establishing rankings for countries based on their innovative capability is R&D intensity, which can be calculated as expenditures on research and development as a percentage of GDP. R&D is considered to influence two complementary processes in enhancing technological capabilities of countries. First, more direct impact, which has also received the most attention in empirical literature, is generating innovation. The second important role of R&D is facilitating the adoption of innovations, which were developed elsewhere, through the development of absorptive capacity. It is said that in order to deliver desired economic results, knowledge always needs to be combined with sufficiently developed absorptive capacity or social capability (Archibugi et al., 2009)(Fagerberg et al., 2007).

The understanding of innovation and innovation processes was previously primarily based on research and development, patent data, and ad hoc surveys. However, there was a significant discourse among researchers and a number of papers have pointed out limitations of those measurements and the neglect of a number of areas, which are important to understand innovation. The most crucial of those critiques are an insufficient coverage of innovation in services, a lack of focus on non-technological innovation, and a deficiency of information on the dynamics of innovation from a systemic point of view (Bloch, 2007).

Economists and scholars frequently measure innovation by other indicators that are alleged to reflect innovation, with the most common choices being R&D expenditures or the numbers of patents applied for or granted. Those measures have also been commonly criticized. Firstly, the R&D expenditure does not necessarily result in the development of new products or processes and some of the innovative firms do not perform R&D. Secondly, a large fraction of innovations are not patented and the importance of patenting also varies between sectors. Patents are also awarded not for innovations but inventions and many inventions never make it to the innovation stage or are not patented, or even patentable. Therefore, using a patent-based measure of innovation is, in many cases, likely to lead to a biased picture. Additionally, measures related to R&D and patents fail to capture innovation occurring through diffusion processes, such as when innovative production equipment or product components are purchased from other firms on the market (Arundel et al., 2007)(Fagerberg, 2013).

An alternative approach is sending questions to firms about their innovation activities, which is the methodology applied in the Community Innovation Survey (CIS) that has been carried out in Europe since 1991. Although it is considered to be more relevant than relying on counting patents, this approach also has its pitfalls. One of them includes differences in sampling techniques, which can make comparisons across countries and over time quite problematic. Moreover, even if the questions in survey are intended to be the same, differences in languages, cultures, contexts etc., can lead firms that are relatively similar to answer differently, complicating the comparisons even further. The interpretation of the results also depends on how innovation was defined in the survey beforehand, and whether or not the firms understand what is meant by innovative activities. The most common definition of innovation is something “new to the firm” rather than to the sector, industry, or to the world, which to some extent could be seen as an imitation than innovation. For those reasons among others, finding a good and reliable innovation metric that can be used to compare countries and regions over time remains a big challenge (Fagerberg, 2013).

Another approach to innovation measurement and analysis, introduced by endogenous growth theorists, highlights the need to introduce technology as an internal factor, which is affecting economic growth. From this perspective, it is argued that there might be increasing returns to investment because of improvements in technology, which in some cases lead to divergence among regions. One of the main concerns in this approach is closely linked to the fact that R&D activities are costly and they require certain critical mass before being capable of generating technological progress, and achieving satisfying economic results. It is particularly important for firms located in regions, which traditionally have lacked a clear-cut scientific and technological strategy. This is a common characteristic of many lagging regions that are also associated with increasingly tighter budgets, which may result in insufficient allocation of resources devoted to R&D. Therefore, it could be concluded, that R&D expenditure might have a small real impact on the economic progress of some regions and would not be a good indicator of growth or innovation (Bilbao-Osorio & Rodríguez-Pose, 2004).

The innovation literature has recognized and highlighted the role of R&D, and the presence of skilled scientists and engineers in order to develop successful innovation in science-based sectors. More recent research within the national innovation systems perspective underlined the importance of other factors for innovative activities, particularly in low and medium technology sectors, where the R&D expenditure frequently plays a secondary role. These factors include other forms of open innovation, interactions with customers and suppliers, and feedback mechanisms from the market (Arundel et al., 2007). There is also a critique about using a single statistical source to measure technological competences of a country, instead of using different sources. For example the data devoted to R&D spending, number of patent applications, or data on high-technology trade can shed light on some specific aspects of technological competences but the analysis will be incomplete. These measures usually underestimate other aspects of knowledge, such as incremental forms of innovation. Therefore, it is said that technological capabilities must be considered in a broader sense, including both the creation of new knowledge and its applications to real economic, and social problems (Archibugi et al., 2009).

Focus on incremental innovation, highlighted in the research by Marshall, on the contrary to Schumpeter’s emphasis on radical innovations, can be seen as an important inspiration for modern innovation studies. It is said that any attempt to link innovation to economic growth and development, needs not only to capture both radical and incremental innovations but also the ongoing processes of imitation and learning (Lundvall, 2017).

Nowadays, innovation capability is seen more as the ability to exploit new technological combinations, also by embracing the notion of incremental innovation. Additionally, the process of innovation has changed significantly over time. Non-R&D innovative expenditure is currently an important component of gaining the benefits of technological innovation. There is an increasing interest in understanding how incremental forms of innovation can impact development, as well as, how innovation takes place in low and middle-income countries. Investments in innovation-related activity have been consistently intensified at the firm, country, and global levels. As a consequence, new innovation actors from both outside high-income economies and non-profit sector emerged. Therefore, the structure of knowledge production activity is currently more complex and geographically dispersed than it has ever been (Global Innovation Index, 2018).

The third edition of Oslo Manual states that innovation is a continuous process, thus firms constantly make changes to products and processes, and acquire new knowledge. Therefore, measuring innovation as a dynamic process rather than a static activity, which makes it even more difficult to capture (Oslo Manual, 2005).

Recent literature on systems of innovations emphasizes the role of learning and building competencies. R&D, as the one important determinant of firms' absorptive capacity, is commonly included in innovation measurement frameworks. But there is also a need to introduce more guidance on measuring human resources and firms activities necessary to build human resource capabilities, for instance trainings, skills development, or job rotations (Bloch, 2007).

Additionally, analysts responsible for collecting innovation data may encounter problems related to demand side and external inputs, which are a part of open innovation approach. It has been argued that users are a crucial source of innovation and creation of new knowledge. The concept of "user-driven innovation" involves the use of knowledge of customer needs and demand in the firm's innovation activities. Another approach called "lead-user innovation" focuses on the participation of lead users in the actual development of new products, where the interaction with customers not only concerns identifying user needs but also involves seeking solutions for the development of new ideas. In this case, users can also potentially be seen as a source of new technological knowledge. Moreover, the benefits of open innovation increase when many new ideas may come from users. On the other hand, it is hard or impossible for firms to predict where new ideas will originate, therefore it causes further issues with the measurement of innovation activities (Bloch, 2007).

An important challenge for innovation system analysis is to understand how different kinds of knowledge are created and used in relation to the process of innovation. Some elements of knowledge are characterized as local and tacit, embodied in people and organizations. Other elements are global, explicit and may easily be transferred from one part of the world to the other. Moreover, different sectors in the economy and in society make use of different mixes of local and global knowledge, therefore there are areas, such as education or business consulting, where it is especially difficult to codify the know-how that is transferred between people (Lundvall, 2017). In addition, some aspects of knowledge can be embodied in the traded commodities, while some others are embodied in the labour force. The limited mobility of labour across national borders can be partly used for explanation why technology is not easily transferred internationally (Lundvall, 2016).

From a cognitive perspective, knowledge is composed both of codified elements (for example scientific and technical literature, patents, manuals, or blueprints), and tacit elements embodied in qualified labour force. Tacit knowledge is considered to be a fundamental element of the innovative process, even though it is much more difficult to measure. Given its tacit nature, it is difficult to quantify, and even if it becomes quantifiable, it also turns into explicit and not longer tacit. Some analyses address this problem by using indicators, which are related to the educational qualifications of employees, under the assumption that education and work experience contain and contribute to development of the tacit knowledge of the labour force (Archibugi et al., 2009).

Many indicators are composed based on "technological capabilities" more than "innovative capabilities" of countries. Innovations are defined as the direct or indirect outcomes of various activities, for example basic research carried out in universities, research and development, production. However, innovations can have different nature, for instance they can be defined as technological and non-technological or tangible and intangible. The advantage of using indicators based upon technological capabilities is that they represent conditions necessary for creation, absorption, and diffusion of technological innovations across the whole economic system. On the contrary, this approach is limited and it cannot fully grasp and illustrate other forms of innovation, especially non-technological ones, as well as, organizational or marketing innovations. Since these forms of innovation are gaining more importance in the competitiveness of countries, innovation measurement frameworks should also address this challenge (Archibugi et al., 2009).

In the measurement framework, which Global Innovation System is based upon, the link between innovation and economic change is central. Oslo Manual (2005) highlights driving forces behind innovation and the importance of not only products and processes, but also marketing, organizational practices, the role of linkages and diffusion, and system perspective on innovation. It is said that through innovation, new knowledge is created and diffused, which allows expanding economy's potential of developing new products and more productive methods of operating. Those improvements depend both on technological and other types of knowledge that are used to develop different types of innovation. These various forms of innovation may differ greatly in terms of their impact on firm's or country's performance and economic change. Therefore, it is important to identify the implementation and impacts of innovations of diverse nature (Oslo Manual, 2005).

The collected data can be organized and divided into two groups: innovation inputs and outputs. Firstly, there are business capabilities, including the knowledge, competencies, and resources that a firm accumulates over time. More specifically, business capabilities involve resources controlled by a firm, general management capabilities of a firm, skills of a workforce and ways how firms manage their human resources, the ability to design, develop, and adopt technological tools and data resources. Those internal factors are of critical importance for analyses of the effects of innovation on firm's performance and on explaining why some firms engage in innovation activities and others do not. Numerous business capabilities may potentially support innovation activities, the development of products or business processes, and the economic impact of these innovations. The newest edition of Oslo Manual also introduces a chapter about a firm's external environment and the associated challenges or opportunities that managers need to consider when making strategic choices. The environment, among other factors, includes legal, regulatory, competitive and economic conditions; firm's customers, competitors, and suppliers; labour markets; and the supply of technological or other types of knowledge of value for innovation (Oslo Manual, 2018).

Furthermore, in the innovation measurement process, it is possible to distinguish innovation outcomes, which refer to the results of innovations, the creation of knowledge, as well as, the impact and diffusion of knowledge. From a firm's perspective, the expected outcomes of innovative activities include an increase in market share, sales, or profits. Increases or decreases in productivity, profits, jobs creation, social and environmental impacts are examples of outcomes of concern for all users of innovation data. At the level of a society, the ultimate impact of innovation is the satisfaction of current or future human needs at either the individual or collective level. Even though measuring to what extent innovation results in social or private outcomes is difficult. Moreover, innovation does not always result in desirable outcomes for all parties. Innovation outcomes might be widely distributed over time, organizations, and individuals. The impacts of innovations can be both measured directly (self-reported impacts), or indirectly through the analysis of data on innovation activities, outputs (such as different types of innovations) or data on internal or external outcomes (such as profits) on different levels of aggregation (Oslo Manual, 2018).

Both innovation inputs and outputs are of an interest to innovation policies. If it comes to inputs, a broad understanding of the distribution of innovation activities across industries, as well as, the role of R&D and non-R&D inputs in the innovation process are particularly important. It is relevant to identify how R&D might be related to other innovation inputs, and to identify other sources of innovations. Even more crucial for policymakers is analyzing the effect of innovation on productivity and employment. It can be analyzed at the national level, for specific regions, or sectors. More thorough information on the conditions for success would help improving policies, which are aimed at achieving the economic and social benefits of innovation. Many policies for supporting innovativeness would also benefit from the identification of main forces which drive firms innovation activities, including market-related impacts or internal organizational changes. It is also significant to identify obstacles to innovation since a substantial proportion of government measures are aimed at overcoming them. The obstacles may include shortages of skills, finance, and appropriation (Oslo Manual, 2005).

5.4 National systems of innovation

The discussions both among innovation researchers and policymakers emphasize the importance of taking a broad perspective on innovation (Oslo Manual, 2005). To conduct an analysis on a country level, it is important to understand the key concepts about innovation systems and how to systematize innovation within national borders. This subchapter introduces innovation systems approach and different ways of aggregating innovations, in order to conduct analysis and comparison between different units. National systems of innovation are the main focus of the theoretical considerations, as the Global Innovation Index is constructed on a country level, introducing similarities and differences between various economies. Additionally, elements influencing the system, as well as, its close environment are important for the creation and diffusion of innovation, thus the description of individual factors that might impact the competitive advantage of countries in the analysis will follow.

In order for different innovation processes to succeed, it requires participation from many different actors, both from private and public sectors, and inputs from a multitude of various sources. When those patterns of interactions or networks achieve a certain degree of stability, it is common to use the term "innovation system" to address them (Fagerberg, 2013). Innovation systems are defined as the determinants of innovation processes, which means all the important economic, social, organizational, political, institutional, and other factors that influence the development, diffusion, and the use of innovations. They consist of different components and relations between them. Among the most important components are organizations and institutions, in the meaning of laws, rules, norms, routines, or other established practices that regulate relations between individuals, groups, or organizations. The fundamental assumption, which highlights the systemic nature of innovation, is that firms do not typically innovate in isolation but rather in collaboration and interdependence with other organizations, including other firms, suppliers, customers, competitors or other entities, such as schools, universities or governments. Those organizations and institutions foster creation and commercialization of knowledge, which causes the emerge of innovations. The function of system of innovation is pursuing innovation processes by development, diffusion, and use of innovations (Edquist, 2010).

Basic idea behind systems of innovation is dated back to 1841 when Friedrich List published a concept of national systems of production. His description of the concept took into account a wide set of national institutions, including those engaged in education and training, together with infrastructure, such as networks for transportation of people and commodities. Some of the crucial ideas that the innovation system concept is built upon, such as vertical interaction and innovation as an interactive process, are also central in the literature on industrial clusters (Lundvall, 2017).

The national industrial systems are divided into a small number of sections, which are defined by the economic function of the output, and by their sectors of use (investment goods/semi-manufactured goods/consumer goods). The process of innovation can be considered as the single most important factor restructuring the system of production by introducing new sectors, breaking down the old linkages or establishing new ones within the system. The interdependency between production and innovation is a point of departure for defining a system of innovation based on the national system of production. Although, it is worth noting that some parts of the production system are more active in terms of innovations, while others are primarily only users of innovations developed by others. Additionally, the national system of production is not assumed to be a closed system. On the contrary, there is some degree and form of openness, which determine the specific dynamics of each national system of production (Lundvall, 2016).

Two main elements essentially form each system of innovation a structure of production and institutional set-up. The first element is associated with the industrial and productive dynamics of a given territory. The institutional set-up is described as a set of widely interpreted socioeconomic and political institutions that support and shape the productive, and technological processes. Since innovation is understood as a complex social process, the existence of specific productive structures and institutional setups are not only the minimum elements which define a system, but they also give each system its unique and unrepeatable character (Borrás, 2004).

The systems perspective on innovation can be applied at various levels of aggregation. Systems can be defined based on technological, industrial or sectoral characteristics but, to a different extent, include other relevant factors, such as institutions (laws, rules, regulations, habits, etc.), political processes, public research infrastructure (research institutes, universities, support from public sources, etc.), financial institutions, skills (labour force), and other components (Fagerberg, 2013)(Edquist, 2010).

The concept of innovation systems has subsequently been applied to other levels of aggregation than the national. The basic motivation for introducing these alternative or additional concepts is that there might be numerous factors specific to regions or sectors. Therefore, it would be relevant to analyze how innovation processes evolve at these different levels. An additional argument for examining innovation systems from other perspectives is that some policies are conducted or targeted at the regional or sectoral level, making other levels of aggregation more useful from a policy perspective (Bloch, 2007).

There are various perspectives that aim to explain the innovation process in relation to specific technologies and sectors. One of the approaches states that innovation systems are unique to technology fields and it is focused on technological systems of innovation, explaining the nature and rate of technological change. It is especially useful in analysis of how new technologies emerge. The 3 other perspectives, including national, sectoral, and regional perspective, are variants of generic "systems of innovation" approach. Sectoral system approach is unique among the different perspectives because it does not define a vertically integrated system as an analytical object. It can be seen as a cross-sectional approach between industrial and innovation economics. The sectoral approach focuses on a group of firms that develop and manufacture various products, belonging to a specific sector, generate and utilize the technologies of this sector. Another approach in the innovation systems literature has the main focus on the spatial level, using national or regional borders to distinguish between different systems. Because the spatial systems are defined on the basis of political and administrative borders, those factors tend to play an important role in analyses based on this perspective, which is proven to be influential among policymakers, especially in Europe (Fagerberg, 2013)(Edquist, 2010)(Lundvall, 2017).

The innovation systems perspective requires interdisciplinary approaches to examine the interdependencies among actors, the uncertainty of outcomes, together with the path-dependent and evolutionary features of systems that are complex and non-linear in their responses to policy interventions. Nations and regions consist of a number of different sectors and technological systems within their borders. Therefore, innovation systems at different levels of aggregation co-evolve and influence each other. A national system of innovation contains firms in many diverse sectors that operate within a common (national) framework (Fagerberg, 2013). Innovation systems can be delineated by industries, technologies, or geography. They are often interrelated, with local systems being linked to national and global systems. Innovation measurement usually involves collecting data at the firm level, which is later aggregated to provide further results at the national or industry level. Systems perspective is used for developing innovation policies in order to coordinate system transformations which serve broad societal objectives (Oslo Manual, 2018).

It is said that the innovation system concept might not only be regarded as a practical tool for designing innovation policies, but also it can be seen as a synthesis of various analytical results produced by scholars working on understanding innovation. National systems of innovation can be used as a tool for analyzing economic development and economic growth. The concept is aimed at explaining how systemic features and different institutional setups link innovation and learning processes to economic performance and to understand better how innovation affects economic development at the country level (Lundvall, 2017).

The idea of a national system of innovation was explored in-depth by the research program pursued by the IKE group at Aalborg University. In the first half of the 1980s, it was referred to as *"the innovative capability of the national system of production"* in several working papers and publications. The term "innovation system" appears for the first time in theories about product innovation and user-producer interaction (Lundvall 1985), although without the adjective "national". The modern version of the full concept and the first definition of national system of innovation (NIS) was introduced by Christopher Freeman in 1987 (Lundvall, 2017)(Edquist, 2010)(Bloch, 2007).

Some influential works, such as Nelson and Winter (1982) and Kline and Rosenberg (1986), highlighted the role of interactions among actors in the development of knowledge and technologies. In fact, Kline and Rosenberg's chain link model of innovation is a base for the conceptual framework for the first and second editions of Oslo Manual. The model presents innovation in terms of the interactions between a firm and other actors, where the completed innovation is the result of a (non-linear) process involving testing, feedback, and subsequent redesign of products or processes (Bloch, 2007).

The initial conceptualizations of national systems of innovation are commonly attributed to 3 founding researchers: Christopher Freeman (1987), Bengt Åke Lundvall (1992), and Richard Nelson (1993). All 3 academics represent different approaches. Their descriptions of the concept itself, as well as, definitions of the terms “innovation” and “system” vary between each other (Rakas & Hain, 2019).

The concept of national systems of innovation introduced by Freeman was described as a network of institutions in both public and private sectors, whose activities and interactions initiate, import, modify, and diffuse new technologies. The main focus of his qualitative analysis was on the ways in which the resources are organized and managed at the firm, industry, and country level, including the organization of R&D and production in firms, relationships between firms, and government's role. Freeman’s concept is seen as the most important in explaining both the emergence and the rate at which the “technological gaps” between countries are closed. Nelson defined the system as a set of institutional actors, whose interactions determine the innovative performance of national firms. The main orientation of his research involves describing the mechanisms and institutions, which support technological advances, relating it to the differences in countries’ economic performances in dimensions like productivity, income growth, export, and import performances. Nelson emphasizes that organizations, described as firms, industrial research laboratories, governmental laboratories, and universities, which promote creation and diffusion of knowledge, are the main sources of innovation. While Nelson puts an emphasis on empirical case studies, where some of them focus on national R&D systems, Lundvall is more theoretically oriented and his research seeks to develop alternative framework to the neo-classical economics, placing at the center of his analysis innovation, interactive learning, and user-producer interactions. Both scholars highlight the importance of organizations supporting R&D. Lundvall defines the structure of production and institutional set-up as the two most important dimensions that characterize a system of innovation. Lundvall’s approach features a broader understanding of the concept and recognizes that organizations are a part of much wider socio-economic system, in which political and cultural influences, together with policies, help to determine the scale, direction, and relative success of all innovative activities within the system. The concept of NIS is viewed as a framework or an analytical and policy tool to link innovation to economic performance at the national level (Edquist, 2010)(Lundvall, 2017)(Rakas & Hain, 2019).

In summary, the differences between these approaches can be attributed to narrower or broader definitions of national system of innovation, main focus of the analysis, and distinguished elements of the system. What is common to all three approaches is the focus on the relationships between the institutions, organizations, as well as, the interactions between them, and learning, innovation, and economic performances at the national level. Different studies on systems of innovation focus on how those interactions operate and on the roles of cultural, organizational, and institutional factors in affecting innovation (Edquist, 2010)(Rakas & Hain, 2019).

The “Aalborg school” places special emphasis on the factors that influence relationships among actors within the innovation system, including all organizations and institutions that affect the production, diffusion, and exploitation of economically useful knowledge. Lundvall’s approach focuses on user-producer linkages and interactive learning as foundations of innovation. Those relationships include norms, habits, and rules. Although there is a number of other factors in national innovation systems that may impact innovation, including laws, regulations, market demand, and institutional organizations (Bloch, 2007).

It has also been argued that if the actors themselves are directly involved in defining the boundaries and the features of the innovation system, it should be explicitly acknowledged in the theoretical definition. Therefore, for a system to be considered as a system, the actors must engage in a process of self-creation, which is based on the social relevance they assign, through their understanding and communication, to its structural elements. The combination of the institutions and the interactions between the actors should generate specific social dynamics, which are intrinsically related to innovation-related activities. This phenomenon has been defined as the "autopoiesis" of social systems, which is taking place when the system itself produces and reproduces the elements of which it consists, through the processes of meaning and communication (Borrás, 2004).

Both Lundvall and Nelson are defining national systems of innovation in terms of determinants or factors influencing innovation processes. However, in their actual definitions of the concept, they distinguish different determinants of innovation. Therefore, both scholars use the same term, even though they propose diverse definitions for it. Moreover, none of the two explanations include consequences of innovations emerging in innovation systems, which are extremely important for socio-economic variables like productivity growth and employment. This highlight the issue of not having a generally accepted definition of national systems of innovation (Edquist, 2010).

Evolutionary approaches, based on the research by Nelson and Winter, view innovation as a path-dependent process, in which knowledge and technology are developed through interactions between various actors, and other factors influencing them. The structure of these interactions affects the future path of economic change. For instance, market demand and the opportunities for commercialization of certain goods influence which products will be developed and which technologies will be successful. Closely linked to the evolutionary approach is an innovation system perspective, as it also studies the influence of broadly defined external institutions on innovative activities of actors, which are a part of the system. The systemic approach emphasizes the importance of the transfer and diffusion of knowledge, skills, ideas, information, and signals of many kinds. The channels and networks through which this information circulates and is being exchanged are embedded in social, political, and cultural background, which guides or constraints innovation activities, and capabilities. Innovation, from a systemic point of view, is viewed as a dynamic process, in which knowledge is accumulated through learning and interacting (Oslo Manual, 2005).

Another, more general, definition of national innovation systems were proposed by Edquist (1997), which describes them as all important social, economic, political, institutional, organizational, and other factors influencing development, diffusion, and use of innovations. He argues that if not all the factors, which influence innovation processes, are included in a definition of a system, it has to be justified which factors should be excluded and what is the reason for that. It is difficult to do so because the exact determinants of innovation are not known systematically and in detail. Additionally, there is not enough information about what are the most important determinants of innovation processes and how different factors affect innovation, therefore they all should be included in the system. Therefore, it might be dangerous to exclude some potential dominants, since they may be considered very important, once the state of art has advanced. This is the reason why this general definition involves relationships among listed factors, as well as, the actions of both public and private entities (Edquist, 2010)(Bloch, 2007).

Other definitions of national system of innovation, adopted further in the literature, are close to the one proposed by Edquist. It was described as a set of distinct institutions, which jointly and individually, contribute to the development and diffusion of new technologies. National system of innovation also provides the framework within which national governments form and implement policies, in order to influence the innovation processes. Therefore, it is a system of interconnected institutions that aim to create, store, and transfer knowledge, skills, and artifacts that define new technologies. The element of nationality occurs not only from the domain of technology policy but also from aspects, such as shared language and culture, which bind the system together. Other important elements are: national focus of other policies, laws, and regulations which condition the innovative environment (Carlsson, 2006).

Since the aim of an analysis of innovation systems is to link innovation to economic development, in order to structure the analysis, it is beneficial to distinguish between the core of the innovation system and the wider setting beforehand. Both should be included in the analysis (Lundvall, 2017). As mentioned already in Edquist's research, a key problem with innovation systems is defining their output - is it innovation in a narrow sense, such as new products or processes, or whether the system also includes the diffusion and use of innovations. The most common approach is using a broad perspective. There are two reasons for that. Firstly, what matters economically is not innovation itself but the changes in production and consumption it causes. Secondly, innovation, diffusion, and their use are interrelated phenomena, with feedbacks back and forth between their different phases. Therefore, applying a narrow perspective, for example not including diffusion and use of innovations, would not only exclude what matters the most economically but also make the innovation dynamics much more difficult to understand. The other issue about innovation systems is defining their boundaries, more precisely deciding what should be included in the system and what should be a part of its environment or other systems. There are also different perspectives on that issue. The first perspective is holistic, in which everything that influences the output of the system is included in it. However, it has been criticized for failing to distinguish between factors that influence each other, and therefore should be seen as a part of the system, and other factors that may have an impact on the system but are not directly affected by it, which should be considered as external parameters (Fagerberg, 2013).

It is also said that there is a need for systematic analysis of different activities in innovation systems. Edquist provided a list of these activities. They are divided into main groups: demand-side factors, competence-building activities, innovation support organizations, and other institutions. The whole list of innovation activities includes: research and development, competence building in the labour force, user feedback and interactions, formation of new product markets, organizations supporting innovation (entrepreneurship, public research, technology transfer, consulting, etc.), both formal and informal networks, institutions (laws and regulations), financial and administrative institutions (Bloch, 2007).

The general institutional environment, as defined in third edition of Oslo Manual, shapes the broad framework within which firms operate. The core components of it include: the basic educational system, which determines educational standards both in among the workforce and domestic consumer market; the university system; science and research base; specialised technical training system; common pools of codified knowledge, including publications and technical, environmental, and management standards; innovation policies together with other policies that may influence innovation; legislative and macroeconomic settings, such as patent laws, tax system, corporate governance rules and policies related to interest and exchange rates, tariffs, and competition on the domestic market; market accessibility - ease of access, market's size, and possibility of establishing close relationships with customers; industry structure and competitive environment, which also consists of analysis of supplier firms in complementary sectors; financial institutions; communications infrastructure, including both transportation and telecommunications network (Oslo Manual, 2005).

It is said that the technological competitiveness of firms is inevitably dependant on national systems of innovation, and likewise, national systems of innovation depend on government policy. This might be one of the premises why is it important to conduct analysis of innovative competitiveness on a country level. The level of business-funded R&D is influenced by different national policies and also by the behaviour of national institutions, such as systems of corporate governance, banks, and stock markets, or agencies funding basic research. Various innovative activities of companies are significantly influenced by their home country's national system of innovation, including: workforce skills, the quality of basic research, system of corporate governance, the degree of competitive rivalry, the price of labour and energy (Carlsson, 2006).

First important factor influencing innovation is the home market itself. Product innovations may be very dependent on the existence of a market for the products. It is even believed that demand influences all innovations, to some extent. Consumers' innovativeness is an important determinant of the rate of diffusion and a tendency among firms to innovate, since it can be observed that the more responsive customers are to newly introduced products, the greater the incentives and return to innovation. Consumers' innovativeness in home markets can also be important for development and testing. Even in cases where domestic markets are not crucial in terms of share of sales revenue, they may be essential for the development and testing of new products. When the value of products or services increases according to the number of others using it, a network effect, also called "network externalities", is present. The term refers to demand, which is affected by the number of users of similar products or the availability of related products. Other factors of importance for establishing demand for certain product groups are standards, platforms, and regulations present on the market (Bloch, 2007).

Countries are also characterized by specific social, political, and economic attributes, which can influence their capacity to transform R&D investment into innovation and economic growth. It has been discovered that different regions have different "social filters" embedded in society. These social filters determine a different capacity for each region to assimilate and transform its own and foreign R&D related innovation into economic activity. As a result, two types of societies are distinguished, described as innovation prone and innovation-averse. The innovation prone societies are the ones, which are capable of transforming a larger share of their own R&D into economic growth and innovations. Conversely, the innovation averse societies cannot manage to transform their own research and development into innovations and economic growth to the same extent. The main components of those social filters responsible for classifying a region are identified as several variables related to local labour markets. The most important ones include: activity rates among workers, unemployment rate, demographics, and educational attainment (Bilbao-Osorio & Rodríguez-Pose, 2004).

The system approach to innovation shifts the focus of policy towards the interplay of institutions and interactive processes of creation, diffusion, and application of knowledge. It also emphasizes the importance of various conditions, regulations, and policies, under which markets operate, as well as, the role of governments in monitoring and adjusting the overall framework. Trends in advanced economies towards greater dependence and increasing need to access knowledge, information, and skills are described with the term "knowledge-based economy". A knowledge-based view is focused on the interactive processes through which the knowledge is created and exchanged (Oslo Manual, 2005).

The importance of skills in the innovation process has been extensively addressed by researchers. It has been highlighted that scientific and technological knowledge are often tacit, and therefore an extensive learning process, based on skills accumulated through years of experience is required in order to codify and transfer this sort of information. Another crucial factor, described as a mechanism to guarantee innovation, is the role of human capital, especially professionals and skilled labor have a vital role in the creation and distribution of innovations. Innovation-averse areas are characterized by including more unskilled labor than the innovation prone regions. Another factor related to human capital, which can influence the innovation process is the situation in the labour market. Low level of activity among workforce and low employment rate are the key characteristics of innovation averse societies. Moreover, having a large, low-educated population is considered a key factor to explain poor innovative performance (Bilbao-Osorio & Rodríguez-Pose, 2004).

The economic structure of a country also plays an important role in the processes of generation and assimilation of innovation. The aspects such as the initial wealth, availability of skills, or the presence of high-technology sectors have an important role in the capacity of a country to transform R&D into innovations. For example, a predominantly agricultural region is characterized as less likely to generate large numbers of patents, as traditional agriculture does not tend to be as innovative as other sectors. On the contrary, certain sub-sectors within the manufacturing and services might be more prone to foster innovations. Particularly economies, which rely on technologically advanced sub-sectors have a greater tendency to note higher rates of innovation (Bilbao-Osorio & Rodríguez-Pose, 2004).

Part of the environment, which might have a positive influence on creation and diffusion of innovations are institutions. The term refers to various rules that influence entrepreneurial actions. They range from law and regulations to informal norms and rules, on which the influence of policy actors might be more indirect. Examples of relevant institutions that influence competitive environment include: intellectual property rights, requirements for setting up or close down businesses, regulations regarding hiring or firing personnel, the prevalence of corruption (or lack of such) and attitudes favoring innovation (for instance social trust, cooperation, etc.). Institutions are mostly considered to be relatively stable but laws and regulations of relevance for business activities do sometimes change. In fact, even attitudes and values in society change in response to technological and economic developments, although the pace of the adjustment is very slow, from one generation to the next (Fagerberg, 2013).

Some researchers argue that formal and informal institutions are equally important in the development, maintenance, and performance of an innovation system. Coriat and Weinstein (2002) have defined “type 1” and “type 2” institutions. The first group includes those that define the rules of the game, which are explicit and imposed universally on all actors through top-down decision-making procedures. The second group includes institutions that emerge spontaneously from interactions between actors, and in a sense, complete type 1 institutions by fine-tuning agents behavior and expectations. Informal institutions are also said to be the ones that assign meaning to the formal institutional set-up (Borrás, 2004).

However, the real effect of both formal and informal institutions in a system lies in their interactions. This effect can be described as the way how each individual actor assigns a meaning to the formal institutions, as well as, the manner in which the design and content of formal institutions respond to these spontaneous informal institutions, causing them to re-adjust themselves accordingly (Borrás, 2004).

The literature on systems of innovation has linked the process of innovation to institutions. The vast majority of literature regarding institutional and evolutionary economics has emphasized the context-bounded nature of technical change, not only in terms of the imperfect information and bounded rationality of the actors, but particularly in terms how relevant the institutional set-up is for innovative performance. Traditionally, institutional economists have identified 3 generic functions of institutions in the economy: reducing uncertainty, managing conflicts and cooperations, and providing incentives. Later on, it has been argued that institutions have at least 5 generic functions in the economy, where the two additional ones include competence-building and defining the limits. In addition to the generic functions, it can be argued that institutions also play more than 10 specific functions in the systems of innovation. The distinction between the generic and specific functions of institutions allows greater analytical precision about the role that they are, or should be, actually playing in the innovation process, compared to the general functions in an economy. The specific functions explicitly address innovative performance rather than the general economic performance. These specific functions include: the production and diffusion of knowledge, appropriation of knowledge, alignment of actors, regulation of labour markets, financing innovation, guidance of innovators, reduction of technological diversity, reduction of risk, and the control of knowledge use (Borrás, 2004).

On the other hand, it should not be assumed that all institutions have positive effects on either economic or innovative performance of a country. In fact, some institutions might even hinder it. Therefore, it is important to assess of the role of institutions for the performance of a specific system and examine them empirically in terms of their ability to fulfill those functions in a satisfactory way (Borrás, 2004).

Another factor, which is necessary for innovation to endure is finance. An access to well-developed capital markets is essential for developing innovative concepts. Some innovative initiatives, particularly from small firms or individual entrepreneurs, or ideas characterized by high uncertainty, technologically or otherwise, can have difficulties in raising the necessary finance in ordinary financial markets, and therefore, the public sector may play an important role in such cases. Moreover, the design of the national tax system and the ease of paying taxes might also be a matter of importance (Fagerberg, 2013).

Another important aspect, related to the process of innovation, is the national government. It has both direct impact through interventions in relation to specific innovations, but also the government can establish standards and regulations to make domestic interactions more efficient (Lundvall, 2016). What policymakers are mostly interested in is the beneficial effects that innovation has, not only for the innovators and/or their immediate environment, but for the society as a whole. Given that such societal benefits in fact exist, a natural question for governments is how innovation can be influenced through the creation, use, and development of various policy instruments. Diverse markets, in reality, are far from theoretical descriptions made by economists, where a self-regulating market would secure a socially optimal allocation of resources. Therefore, so-called “market failures” provide justification for market interventions or for introducing policy instruments which aim at increasing investments in science or research, towards the socially optimal level (Fagerberg, 2013).

The concept of national innovation systems has become very widely used perspective both in studying innovation, and in analyzing science and technology (S&T) policy. In terms of the share of overall S&T policy funding, research and development support is by far the major instrument, which implies that science and technology policy still generally follows a linear model of innovation. However, there has been a number of implemented policy initiatives, which focus on interactions and improving firms' and other organizations' framework conditions, such as those involving industry–science relationships, technology transfer, and innovation incubators. These policy measures are relatively small when it comes to expenditures, yet they provide an indication of the growing interest on viewing innovation from a systems perspective among policy makers (Bloch, 2007).

Government may interfere, either directly or indirectly, in relation to both the establishment and restructuring of patterns of user-producer relationships. In some important instances, the state intervenes directly in the network and supports existing user-producer relationships. It may also support the establishment of specific organizations necessary for intermediating between groups of users and groups of producers, to pool information, and thereby, stimulate the processes of production and diffusion of innovations. In periods characterized by incremental innovations and gradual technological change, a national government can sustain national and international user-producer linkages. In periods characterized by radical innovations and a shift in technological paradigm, the role of government becomes more important and complex. Then, a transformation of existing network of user-producer relationships is needed. One of the difficult tasks for governments is stimulating the renewal or severance of well-established user-producer relationships, and the establishment of new relationships, when on the organized markets at the national level there is inertia supported by the political power of strong interest groups, closely associated with the dominant structure (Lundvall, 2016).

The role of government is also central for implementing strong incentives and regulations to drive the transition. Governments often take risks by promoting mechanisms that stimulate investments and the diffusion of technologies with disruptive potential. The government can help by supporting the creation of markets for innovative solutions, for example by subsidizing, using public procurement proactively to foster innovation, or changing standards and regulations, so that demand for new, innovative solutions would be strengthened. Different interventions of governments might take different forms, for instance financing organizations dedicated to doing science, such as universities and research institutes; subsidizing research activities in private firms; or attempting to change the rules of the game by strengthening intellectual property rights (Fagerberg, 2013).

The main users of innovation data is the policy community, consisting of policymakers and policy analysts. One of the important functions of innovation data is to provide an informed basis for public policy decisions through benchmarking indicators, and conducting research using collected data on innovation activities. There is a need for coherent policies across multiple government agencies, which are required to mobilize the transformational power of innovation, in order to achieve key policy objectives. A major interest to the policy community is understanding the effects of innovation policies on the innovation activities of different organizations, especially firms. Innovation policies are intended, as their primary or secondary objective, to influence the nature and extent of innovation in an economy. The implementation of innovation policies or practices can be complex. It can be influenced not only by the intention of authorizing legislation, but also by the actual use at various organizational and jurisdictional levels. Innovation policies require coordination and institutional arrangements that often extend beyond science and research ministries to the whole government (Oslo Manual, 2018).

There is a broad range of processes that have impact on the technological dynamics of nations, these processes are also influenced by a large number of policies and actors. However, most of these policies are not defined as "innovation policies" and have not traditionally been regarded as such either. Nevertheless, it is possible that effects of other policies on innovation are much more important than those, which were specifically designed to influence innovation, and are narrowly defined as such. From an innovation system perspective, what matters is not just the name of a policy, but its impact. This carries important implications for how the analysis of the relationship between different regulations and technological dynamics should be conducted. The innovation system perspective leads to a much broader perspective on policy, more specifically on the number and type of policies to take into account, than what has been common among analysts. Moreover, since the processes and innovation activities, which are influenced by these policies are seen as interdependent, as a consequence the effect of a specific policy cannot be assessed in isolation, independently of other relevant policies. Thus, the innovation system perspective leads to a holistic perspective on policy, which has been described by the term "holistic innovation policy" (Fagerberg, 2013).

This holism makes it more challenging for policymakers. Firstly, calculating the total effects of a broad set of interrelated policies (processes) requires larger and more sophisticated analytical capacity in public administration. Applying the innovation system perspective to policy analysis would mean that policymakers from different domains (ministries, sectors, administrative levels etc.) would have to work together and coordinate their activities, which makes the use of holistic perspective even more complicated. This challenge is not only limited to innovation policies and it is difficult to achieve, as this approach tends to collide with the already established structures, practices, and routines in public administration. Therefore, successfully applying the innovation system approach might require the development of new "systemic instruments" that would help to facilitate the creation, adaptation, and coordination of policies, so that the different policies support rather than contradict each other (Fagerberg, 2013).

Innovation systems approach can be seen as one of the specializations of the evolutionary agenda, putting an emphasis on informing science and technology policy. Since such policy has been largely formulated at the national level, the development of the National Innovation System concept was the initial focus. Moreover, many contributions to the field of systems of innovation research, and research on innovation, R&D and technological change in general, were enabled and driven by the public policy needs. It has been noted that external engagements with governmental and intergovernmental organizations were instrumental for the formation of the field of innovation systems, especially on the national level, and continue to shape its development. For instance, OECD undertook substantial efforts to promote and diffuse the systemic approach to innovation among policymakers across the globe (Rakas & Hain, 2019). Therefore, it is not only the theoretical concept that shapes policies but the relationship is mutual and governmental organizations also influence the development of systems of innovation studies.

National systems of innovation can differ greatly between countries, and policies that function well in one context might be totally inadequate in another. As a consequence adopting an innovation system approach leads to a scepticism towards policy advice that advocates the same solution everywhere, independent of cultural and contextual differences (Fagerberg, 2013).

Freeman argues that in order to increase the internationalization of innovative activity, national and regional systems of innovation remain essential. Their crucial role derives from the networks of relationships within them that are necessary for firms to innovate. While external international connections are becoming increasingly important, the influence of national education systems, technical and scientific institutions, policies, industrial relations, traditions, culture, and many other national institutions is still fundamental. In many ways, those factors are what makes each system unique. They represent historical events, their influence and change, only very gradually, and by that create strong path dependence. Nelson additionally remarks that there are both similarities and differences among countries in institutional arrangements, which remain persistent over time, and therefore the distinctive national character of innovation systems is likely to stay unchanged (Carlsson, 2006).

It is also believed that without a broad definition of the national innovation system, which includes individual, organizational, and inter-organizational learning, it is impossible to establish the link between innovation to economic growth. There should be even more careful consideration when the attention is given not only to the science infrastructure but also to institutions or organizations, which support building competences in labour markets, education, and working life. This is especially important in the current era of the learning economy that becomes more globalized (Lundvall, 2017). In such a learning economy, the speed of the innovation process is said to be the critical factor in economic performance (Arundel et al., 2007).

The innovation system perspective is considered not only as a tool to explain innovation but it was also developed as an alternative analytical framework to challenge standard economics when it comes to explaining competitiveness, economic growth, and development (Lundvall, 2017).

5.5 Competitiveness

Over the years, there has been growing interest in analyzing competitiveness of countries (Fagerberg et al., 2007). However, in order to compare different economies, it is important to establish an understanding in which way the competitiveness between countries is interpreted by the researchers and what are the factors influencing it. This subchapter includes definitions of competitiveness on a macroeconomic level, concepts of cross-country convergence or divergence, and the description of the process of catching up by the countries that are falling behind vis-a-vis frontiers. Those concepts will be helpful in interpretation why there are differences between countries, the dynamics of changes in competitive environment, and what could be the factors that influenced the establishment of leadership by some economies.

When competitiveness is considered at the country level, it is described as the way in which the pattern of international trade changes over time to reflect evolving patterns of capabilities, and hence competitive advantage, rather than the established patterns of comparative advantage, which is the usual focus of trade theory (Cantwell, 2005). It is said the competitiveness of countries does not represent the absolute performance of a single economy but rather how well a country performs in relation to others. It is usually measured by both the economic well being of citizens, most commonly measured by GDP per capita, and the trade performance of this country (Fagerberg et al., 2007).

Competitiveness is also defined as the possession of the capabilities needed for sustained economic growth in an internationally competitive environment, where there are other countries, clusters, or individual firms, which have an equivalent but differentiated set of capabilities of their own. The term competitiveness can be used to imply a continuing rise in the living standards of the individuals that are members of a social group with the required capabilities. In this particular context, to imply a sustained increase in the living standards of the citizens of a country that has some competitive advantage in world markets (Cantwell, 2005).

Traditionally, there has been a tendency among many economists to examine competitiveness based on overly simplified models. Standard growth models were developed under the assumption of a closed economy, in which new technologies are equally accessible for all actors, sectors, regions, and nations, being perfectly free and mobile. Additionally, standard foreign trade theory assumes that labour and capital are perfectly immobile, and commodities to be perfectly mobile across national borders. Those assumptions are at odds with what can be observed in the real world, where it is common that some countries establish themselves as technological leaders, in general or in specific technologies, while others tend to lag behind. In accordance with theuser-producer approach, geographical and cultural distance is a factor that can disrupt the interaction between users and producers, which may contribute to an explanation why different national systems display different patterns of development (Lundvall, 2016).

In earlier times, it was believed that technology was mostly embodied in persons, therefore migration of skilled workers was a necessary prerequisite for the diffusion of technology across different locations. However, with the rise of the “machine technology”, this logic had changed. Conversely to the conditions that had prevailed previously, it has been argued that the new type of knowledge and technology is no longer tacit and embodied in persons, but it became more codified and easily transmittable. Thus, catch-up should be relatively easy, since the late-comers could take over the “ready-made” new technology, without having to share the costs of its development (Fagerberg & Godinho, 2004).

This description assumes that technology is easily available and transferable, not very demanding in terms of skills or infrastructure, and that market forces are capable of taking care of the necessary coordination, without any large-scale involvement of external change agents. However, there is a contrasting theory, assuming that technology transfer is so demanding in terms of skills and needed infrastructure that market forces, if left alone, are considered unlikely to lead to success. Therefore, some degree of active intervention in markets by outsiders, such as private organizations or government, is consequently deemed necessary. In this view in order to succeed, catching-up countries have to build up new institutional instruments for which there previously was little or no counterpart in the established industrial country. The purpose of these institutional instruments is to mobilize resources needed for the necessary changes that modern technology requires (Fagerberg & Godinho, 2004).

Neo-Schumpeterian economists have particularly emphasized the connection between structural change and growth through innovation. Their approaches to international competitiveness focus on forging technological competitiveness, which implies a sustainable increase in the share of world trade for those countries, whose innovative efforts are the most successful. However, in the Schumpeterian perspective competition results in establishing new areas of value creation, so innovations expand the overall importance of world trade and the world market. The actors that contribute most to this process of expansion will see their shares rise because they are responsible for value creation. Additionally, the very nature and purpose of innovative activities is to disturb and generate income, in an experimental and non-equilibrium manner. Nowadays, innovation and competitiveness are distinguished with an emphasis on longer-term technological “non-price” competition, as opposed to previously studied shorter-term price and cost-based competitiveness (Cantwell, 2005).

Competitiveness can be analyzed from different perspectives. One of them, deriving from the perfect competition, treats technology as a public good. However, it can be assumed that different technologies are cumulative and context-dependent, which prevents the economic benefits of innovation to spread automatically. But the diffusion of technology from the developed part of the world can serve as a powerful factor behind competitiveness and growth in the low-income economies. Moreover, the technological gap between a frontier and a latecomer provides an opportunity to imitate more advanced technologies elsewhere. But there are some requirements for countries that need to be met in order to get the most of such opportunities, which are called “capacity competitiveness”. They are additional to 3 other types of competitiveness: price, technology, and demand (Fagerberg et al., 2007).

The first type of competitiveness, reflecting the traditional approaches focused on price/costs comparisons, is price competitiveness. It is aimed at emphasizing the potentially damaging effects of excessive growth of costs or prices on the economy (Fagerberg et al., 2007).

Another kind of competitiveness, based on technology, refers to the ability to compete successfully in different markets with new goods and services. It is related to the innovativeness of a country, measured by different data sources reflecting various aspects of the phenomenon. The indicators include for instance R&D expenditures, patent statistics, and ICT infrastructure, measured by the number of personal computers or phone landlines etc. (Fagerberg et al., 2007).

The next type of competitiveness is capacity competitiveness. Sometimes, it might not be easy to clearly distinguish this type of competitiveness from technology competitiveness, since some resources devoted to developing new goods and services can also be beneficial for the abilities to exploit economically such innovations, and vice versa. The focus of capacity competitiveness is on the capabilities which are important for the capacity to exploit technological opportunities. One of the terms used to describe such phenomenon is “social capabilities”, which consist of 3 general factors of relevance: technical/organizational competence (level of education), availability and quality of financial institutions or markets, and quality or efficiency of governance (Fagerberg et al., 2007).

The last category of competitiveness is related to world demand. It has been distinguished based on the premise that the relationship between a country's production, or trade, and the composition of world demand might be important for competitiveness between countries. It is said that demand is not likely to grow at the same pace for all products, therefore products based on important innovations in not too distant past are particularly likely to experience high growth in demand. This might affect countries differently, depending on their specialization patterns (Fagerberg et al., 2007).

There have been also discussions on competitiveness at the industry level, which is described as the nexus of relationships between firms and their environment. Different groups of firms that form the major players in an international industry are referred to by the term of industrial leadership. Such leadership might be achieved because of the national or regional environment in which firms operate, or the institutions that are industry-specific, and not just because of the factors that are purely internal to the firms. Based on the case studies of the evolution of national industryies, it has been concluded that competitiveness derives from the contributions of each, and the interactions between firms, regions, and countries, as well as, the sectoral support systems that connect those different levels of analysis (Cantwell, 2005).

The factors that are thought to influence competitiveness can be grouped into resources or capabilities, institutions, markets or demand conditions, and inter-company networks. Those elements of the environment were described already in the context of national systems of innovation. The relationships that exist between the firms that develop the technological capabilities and various institutions vary from one country to another. But they tend to be particularly different between countries that belong to an already leading industrialized group, and those that attempt to catch up with them. It has been noted that there have been numerous cases, in which governments in catching up economies have actively contributed to the fostering of capabilities in local infant industries and in domestic firms, partly through means of domestic protectionism. There are also examples of economies that have taken advantage of various aspects of trade liberalization instead. Of course, when the domestic firms caught up and the countries themselves sometimes forged ahead, and became innovative leaders, the institutional structures of catching up economies changed considerably and the protectionist measures were largely reversed. This is the most vivid illustration that the development of technological capabilities in firms, and the character of various institutions that support these competitive efforts in the wider society, tend to co-evolve with one another through a process of continuous interactions (Cantwell, 2005).

The history of capitalism since the industrial revolution shows increasing differences in productivity and between living conditions across different parts of the globe. 250 years ago the difference in income or productivity per capita between the richest and poorest country in the world was approximately 5:1, nowadays this difference has increased to 400:1. However, despite this long run trends towards divergence in productivity and income, there are many examples of initially backward countries that, at different times, have managed to resist this trend by narrowing the gap in productivity and income between themselves and the frontier countries, by "catching up". Researchers tried to address the questions about: how did they do it and what was the role of innovation and diffusion in the process. The "catch-up" should be considered in distinction from the discussion of convergence, although the two issues overlap partially. The process of catching up relates to the ability of a single country to narrowing the gap in productivity and income vis-à-vis a leader country. Convergence, on the other hand, relates to a trend towards the reduction of the overall differences in productivity and income in the world as a whole. Obviously, if all countries below the frontier caught up, convergence would necessarily follow. However, if only some countries catch up and perhaps forge ahead, while others fall behind, the outcome with respect to convergence is not clear. The empirical studies show that such convergence is only limited to groups of countries in specific time periods. But in order to explain such differences in the conditions for catch-up, it is not enough to rely on general mechanisms, also a historical perspective is necessary (Fagerberg & Godinho, 2004).

The substantial interest in international competitiveness and variations in growth rates has created a new branch of literature addressing cross-country convergence or catching up versus divergence or falling behind. One of the conclusions researchers came up with is that whether an analyst observes convergence or divergence depends upon the studied period and selected countries. Nonetheless, the overall trends in variance between countries are not the most important issue, but rather the understanding how firms in one region had the capabilities to catch up in the given period, while economies in other part of the world did not. The concepts of a techno-socio-economic paradigm, or of evolution in the institutional characteristics of capitalism, may be useful as means of explaining occasional shifts in technological leadership or longer-term competitiveness, and the direction of those shifts. The emphasis has been placed on the role of structural change in growth, particularly during periods of paradigm change, when the predominant characteristics of innovation undergo transformation, which helps to explain the existence of windows of opportunity through which the catching up of some countries may be especially substantial. During these periods leaders might have particular difficulty in adjusting to the new conditions, since they have become locked into the types of innovations favored under the previous paradigm. Meanwhile others, that fall behind initially, may find that their different institutions and methods of social organization are in fact well suited to adapting, and to promote the kinds of structural change in which currently the greatest opportunities for innovation are placed (Cantwell, 2005).

There is also a research on catch-up, on the macro-level, which addresses questions about to what extent catch-up or convergence has actually occurred, for whom, and how this may be explained. An important finding in this literature is that the long term trend since the British industrial revolution shows divergence, not convergence, among capitalist economies. It has also been discovered that these trends differ greatly between time periods (Fagerberg & Godinho, 2004).

The literature on catch-up processes in Europe has led to a strong focus on the relationship between catch-up, institutional instruments and policy. However, the discussion of catch-up and policy has older roots. Already 200 years ago, researchers in the United States based on the authority of Adam Smith argued that the best strategy would be practicing free trade, refraining from governmental intervention in economic affairs, and stick to America's acquired advantage in agriculture. Others advised an industrialization policy based on so-called "infant industry protection", which is related to trade protectionism. There has also been research focusing on the role of banks in industrialization, which was seen as an attempt to form a more general theory about catch-up, emphasizing certain requirements that need to be met for successful catch-up, and institutional responses or catch-up strategies. This view put the emphasis on the advantages of targeting rapidly growing, technologically advanced industries. However, it should be pointed out that the generalization was based on historical evidence. Therefore, it is not obvious that these recommendations would be equally relevant for other time periods or technologies (Fagerberg & Godinho, 2004).

However, what is a suitable policy nowadays, depends not only on the characteristics of the policies that worked well in the past, but also on the current economic, technological, institutional, and social context, which might be quite different from those of previous periods. It is said that the conditions for catch-up have become more demanding over time, with greater demands on the technological capabilities and innovative efforts of countries striving to narrow the gap between the leaders. While in the 1960s and 1970s the main factors supporting catch-up were capital accumulation and a sufficient manufacturing base, in the 1980s and 1990s it has changed and accumulation of technological capabilities, and specialization in services became more relevant. These findings indicate that what has happened cannot be explained solely by changes in institutions and policies, but also a shift in the underlying technological conditions needs to be considered (Fagerberg & Godinho, 2004).

It has been suggested by Fagerberg and Verspagen that the observed shift in the conditions for catch-up might be a reflection of the radical technological change in the last decades, emphasizing that ICT-based solutions have been substituting earlier mechanical and electromechanical ones, and consequently the derived change in the demand for skills and infrastructure. This concept is related to technological congruence. Thus, it can be assumed that, compared to the situation three or four decades ago, the progressive technologies have become less compatible with economic conditions (particularly, the skill-base and R&D infrastructure) that dominate in many developing countries. In fact, nowadays only the countries that have massively invested in the formation of skills and R&D infrastructure seem to be able to catch-up, while those that have not, fall further behind (Fagerberg & Godinho, 2004).

The differences in performance across countries over time may, to some extent, be explained with the help of two concepts: technological congruence and social capabilities. The first concept refers to the degree to which the leading and follower country characteristics are compatible in areas such as market size, factor supply etc. The second concept of social capabilities focuses on various efforts and capabilities that developing countries have to establish in order to catch up. Those include: improving education, infrastructure, and more generally, technological capabilities (Fagerberg & Godinho, 2004).

What is the most noticeable in all the instances of successful catching up, is that the various trade policies of governments were just part of a much wider package of support for the longer term development of capabilities in indigenous firms. Since the end of the 19th century and the emergence of science-based industries, one of the examples of fostering innovativeness was investing in science and higher education, the training of engineers, and more widely in the gathering of various skills. The need for an infrastructure that suitably supports relevant education, skill formation, and training has become critical to the competitiveness of industries, and became even more important in the modern techno-socio-economic paradigm, associated with digitalization and information processing (Cantwell, 2005).

For firms, in order to create capabilities, it is required to undertake costly and difficult internal learning processes. But these processes depend on having suitable organizational and technical skills in both management and workforce, on which they rely. Therefore, the composition of skills in the workforce within the domestic firms is critical to the success or failure of countries that are trying to catch up. But it also becomes a major influence upon the fields, in which any national group of firms has a comparative advantage in innovation and capability creation. In addition, different investments and commitments to training that are made by firms, the professional associations they form, and the pressures they place upon governments and other institutions, imply a process of co-evolution between firms and their environment (Cantwell, 2005).

In summary, competitiveness originates from the creation of locally differentiated capabilities that are needed to sustain growth in an internationally competitive environment. Those capabilities are created through innovation. Because capabilities are varied and differentiated and the creative learning processes for generating them are open-ended, and generally allow multiple potential paths to success, as a consequence a range of different actors may improve their competitiveness together. Innovation is considered to be a positive sum game, which often consists of joint efforts of many to develop new fields of value creation, where, on average, the complementarities or spillovers between innovators tend to outweigh negative feedback or substitution effects, even if some actors lose ground or fail. To conclude, efforts to promote competitiveness through innovation can hardly ever be understood in isolation from what other actors are achieving at the same time. This applies on a country level, within national groups of firms in an industry, sub-national regions, or on a level of individual firms. It is also worth emphasizing that the degree of interaction between innovators in search of competitiveness has been substantially rising historically, and has reached new heights in recent years (Cantwell, 2005).

6. Analysis

The analysis will be based on the Global Innovation Index itself and the key findings presented in the most recent edition of the report, from the year 2018. The findings will be interpreted based on understanding of theories included in the previous chapter and will help to understand the second part of analysis, based on the quantitative study of the dataset, on which the Global Innovation Index is based upon. The quantitative analysis will include Exploratory Data Analysis and Principal Component Analysis to find some patterns in data, and hierarchical clustering, in order to group countries based on their characteristics.

6.1 Theoretical Analysis

The first question that may be raised, while conducting analysis on a macroeconomic level, is whether it is possible to compare such diverse national systems of innovation. A favorable premise for the comparison is that there are more similarities between economies nowadays than two decades ago. Until 1980s the economies were distinguished between the advanced capitalist nations, called “first world” countries, the planned economies, called “second world”, and the “third world” countries, which described less developed nations. They were all characterized by different technological capabilities. For instance, planned economies did not have a business sector to develop innovations in a competitive base. Additionally, the lack of proper intellectual property rights system did not allow to use patents as a technological indicator for those countries. At the same time, planned economies were characterized by a combination of high investments in R&D, as well as, well-educated population and a high level of qualifications among workforce. This, together with indicators based on years of education or number of engineers, made planned economies closer to advanced nations, even though other measurements, for example number of patents, placed them closer to underdeveloped countries. As a result, it was difficult to rank countries belonging to different groups. The disappearance of planned economies has made it easier to rank and compare countries, since all of them can be assessed according to the same criteria (Archibugi et al., 2009).

It is said that innovation can be a major competitive advantage for any developed economy. However, as it was stated before in the theoretical considerations of this project, innovation is a complex and difficult outcome to measure. Additionally, there are many different variables that contribute to innovative performance of each economy at a national level. Research and development (R&D) expenditure is certainly considered as one of these factors, historically one of the most important ones used to measure innovation. Even though in recent years scholars have stated that it does not always directly correlate with innovation outcomes, it can still represent time, capital, and efforts that have been put into researching and designing the products of the future (Desjardins, 2018).

Over the last three decades, there have been positive shifts within the global landscape of investment in science and technology, as well as, in education and human capital. Nowadays, it is no longer only a few high-income economies, such as the United States, Japan, and certain European countries, that carry out research and development. R&D is now a common pursuit or a serious policy ambition in most economies. The worldwide estimated total of R&D expenditures has continuously been rising, more than doubling over between 1996 and 2016. Similarly, the R&D intensity, defined as R&D expenditures divided by GDP, has also intensified over recent years, rising on average, from 1.5%-1.7% between 2000-2016. However, R&D is still highly concentrated in high-income economies (Global Innovation Index, 2018).

In the first phase of the analysis, the ranking from Global Innovation Index will be compared with the national research and development (R&D) expenditure, which was recently presented by the World Economic Forum. The statistics for the visualizations are based on data collected through a series of smaller regional surveys by UNESCO Institute for Statistics. The values were adjusted to reflect purchasing power parity (PPP$), which makes it possible to compare expenditures between economies. The presented numbers include governmental, private, academic, and non-profit investments in research and development, showing a complete overview of the R&D market across the world (Amoros, 2018).

Rank	Country	R&D Spending (PPP)	Global share (%)
#1	United States	$476.5 billion	26.4%
#2	China	$370.6 billion	20.6%
#3	Japan	$170.5 billion	9.5%
#4	Germany	$109.8 billion	6.1%
#5	South Korea	$73.2 billion	4.1%
#6	France	$60.8 billion	3.4%
#7	India	$48.1 billion	2.7%
#8	United Kingdom	$44.2 billion	2.5%
#9	Brazil	$42.1 billion	2.3%
#10	Russia	$39.8 billion	2.2%
#11	Italy	$29.6 billion	1.6%
#12	Canada	$27.6 billion	1.5%
#13	Australia	$23.1 billion	1.3%
#14	Spain	$19.3 billion	1.1%
#15	Netherlands	$16.5 billion	0.9%
	All other countries	**$249.8 billion**	**13.9%**

Figure 3: R&D spending (Amoros, 2018)

The table above presents 15 countries that are characterized by the heaviest R&D spending in the world. As can be seen, they account for 86% of the entire world's outlay. The United States is far beyond the rest of the globe, contributing more than $100 billion more than China on the second place of the ranking. In fact, almost 27% of the global share that is accounted for the leader's expenditures is significantly more than the bottom 100 countries combined. It is stated that investment in R&D can predict how countries will develop in a long term (Amoros, 2018).

It is interesting to compare those figures with the innovative performance in the Global Innovation Index, which is based on many more indicators that take into account different perspectives on innovation. It is clear, that when other factors than only R&D expenditures are taken into consideration, the ranking does not look the same. Even the top 3 economies with the biggest investments in R&D do not remain in top 3 positions in the overall performance in GII.

The United States is placed on the sixth rank both in the overall performance and innovation inputs, taking seventh place in innovation outputs. Worth noting that the U.S. is not shown at all in the top ten economies listed based on innovation efficiency ratio (calculated to analyze how much innovation outputs a given country is getting for its inputs), which can be interpreted in the way that there are not so many returns on R&D investments. The other explanation for the score could be that innovation outcomes are difficult to measure, therefore some of them are not represented in indicators included in collected data. The score of efficiency ratio for the United States is 0.76

United States is ranked first for the market sophistication pillar, indicating that it is a good place to start a business, moreover it leads the ranking for state of cluster development with all the high-tech firms located in Silicon Valley. It is also ranked as 6th in domestic patent applications, mainly driven by big high-tech companies protecting their inventions. The country has first place in both scores by top 3 universities and R&D expenditures of the top 3 firms. However, these indicators are quite limited, and cannot assess the average level of all firms or regions within a whole economy.

When UNESCO’s data, which visualizations of the biggest R&D spenders was based upon, was analyzed further by the researchers working on the index, it was revealed that whereas the U.S. employs 4,295 researchers per million inhabitants, China has only 1,096. But it should not be forgotten that China has many more inhabitants living in the country than the U.S., therefore the numbers cannot be compared directly (Amoros, 2018). In fact, when the information included in the Global Innovation Index (2018) is analyzed, it becomes clear that the total number of researchers for China is the highest in the world. Therefore, it is very important while making international comparisons to use the units that can be directly compared, and not ignoring other factors that might have an influence on an indicator, like the big difference in populations of compared economies in this case.

As it was stated in theoretical considerations of this project, countries are characterized by specific social, political, and economic attributes, which influence their capacity to assimilate and transform domestic and foreign R&D into innovations and economic growth. Human capital, especially educated and skilled labor has a crucial role in the creation and distribution of innovations, and characterizes innovation prone societies (Bilbao-Osorio & Rodríguez-Pose, 2004). Moreover, it has been argued that formal and informal institutions are equally important in the development, maintenance, and performance of an innovation system (Borrás, 2004). The measures related to the number of researchers or tertiary enrolment for China are quite high and can be linked to national characteristics related to demographics, simply assuming the causality that if the population is bigger, the number of students or researchers will be higher as well. Also knowing the social norms in China, it becomes clear that education is an important part of raising future generations. Therefore, both formal and informal institutional set-up in China favors innovation. The advantageous competitive position of China is also caused by the fact that it is a low-cost country and many multinational enterprises, especially in high technology, locate their production centers in China. This could also be one of the reasons why the R&D expenditure in China is that high.

Indicator for research talent in business enterprises is higher for the U.S. (71.1% 4th rank) than for China (61.9% 9th rank), referring to researchers engaged in the conception or creation of new knowledge, products, processes, methods, and systems, in firms, organizations, and institutions, primarily focused on production of goods or services. U.S. also leads the ranking for citable documents H index, while China is 14th. Yet, both countries perform badly in the number of scientific & technical articles - U.S. on 43rd place, and China on the 42nd. This might indicate that both countries put more emphasis on using and commercializing the knowledge or technologies in private sector, than producing scientific outputs. Higher level of citations for the U.S. can also be related to the linguistic advantage and good reputation of American universities.

China, placed second based only on R&D expenditures, is positioned as the most innovative upper-middle income economy in all sub-indices within Global Innovation Index. It is also ranked on the 17th position in the overall ranking. One of the reasons for exceptionally good scores in innovation outputs is that the growth in intellectual property (IP) filings in recent years have been mostly driven by China. The main explanation for that is the regulatory environment and strong protectionism embedded in China’s social and political norms (Global Innovation Index, 2018). But again, the most interesting indicator to look at is the innovation efficiency ratio. In China’s case, the country is placed on the third position in the whole ranking with the impressive overall ratio of

0.92, which is mostly driven by protecting IPs, trademarks, industrial designs, and exports of both high-tech and creative goods.

Japan, which is the third biggest R&D spender, is not even presented in the top ten countries in Global Innovation Index, as its overall position is 13th. It is important to look at Japan's efficiency ratio to get more insights into how R&D investments are linked to the measured innovation outcomes. The third biggest investor in R&D does not perform well in this figure with the score of 0.68, placing the economy on the 44th rank in the world. The country's ranks for all input pillars vary from 8th for institution to 16th for human capital. It also scores well (12th rank) for knowledge & technology outputs, led by first places for patent applications both domestic and international.

The literature on catch-up processes led to a strong focus on the relationship between catch-up and the important role played by institutional instruments, and policy. Different policy instruments have been used to support the growth of new industries or trade protection. Japan and other Asian countries applied the interventionist policies during their catch-up, and many developed countries still do that (Fagerberg & Godinho, 2004). So, similarly to China, Japan has long history of protectionism, which may explain the economy's leadership in patents applications.

Japan indeed has high scores related to R&D, leading the ranking for GERD financed by businesses, and being on the 3rd place for GERD performed by businesses, indicating that R&D is crucial for private sector. However, the country still underperforms in terms of innovation outputs and the main reason behind that is creative outputs. The country has low scores for all indicators in this pillar, except for 6th place for entertainment & media market consumers and advertisements spending. There might be 2 reasons for that: firstly, Japan is focused more on technology than production, which would explain much lower scores in trademarks or industrial designs than China, and low level of cultural & creative goods or services exports, as it is not the main sector in Japan. Another explanation could be the data collection itself, for instance 2 indicators regarding ICTs & business models are based on the surveys, thus the results are subjective.

Based on data from various composite indicators, it has been tested whether a single measure such as R&D intensity can provide similar rankings to composite indicators, and if the R&D intensity can be a proxy of innovative capabilities, capturing similar overview of reality presented in various indicators. The main advantage of composite indicators is said to be the ability to synthesize all the collected information about technological capabilities of a country in a single number. The hypothesis to test was formed on the premise that intra-country comparisons are based on a single number, which comes from some synthetic indicator, independently from the fact that those indicators are based on number of different statistics. Taking this into consideration, if R&D-based rankings turned out to be identical or very close to synthetic indicators, there would not be any point in spending time on constructing them. Therefore, trying to answer the question whether the one-dimensional R&D measure would be able to synthesize country's innovative performance, as well as, the multidimensional composite indicators seemed to be reasonable (Archibugi et al., 2009).

The hypothesis was tested on two separate groups of countries - first one with the highest rates of R&D intensity, and second one with much lower scores. The first group of observations, including more developed economies, was not only more heterogeneous but also the data regarding competitive capabilities and innovative activities for those countries is more reliable and complete, thus it allows to assess the robustness of results for the first subgroup. Researchers concluded that when homogenous groups of countries are taken into account, R&D intensity is less capable of grasping differences between them, as non-R&D factors play an important role in differentiating national performances and paths of innovation. As a result, rankings based only on R&D intensity would only be able to explain differences in innovative capabilities in the case of large and heterogeneous group of countries. Another finding of the study was that correlations rates between scores of composite indicators and R&D intensity drop significantly for the group of countries with low R&D intensity and it is problematic to distinguish characteristics of innovative activities of countries that do not perform, or perform very limited, R&D. Thus, R&D intensity would not be able to capture the reality of synthetic indicators, based on many different variables. Based on the comparison of correlations, it was concluded that composite indicators are more useful, especially in comparing homogeneous groups of countries and highlighting what differentiates analyzed economies (Archibugi et al., 2009).

The R&D intensity itself is not capable of grasping differences in innovative capabilities between all the countries, so it is better to use composite indicators, which explain a wide range of phenomena crucial to understanding innovation, especially to compare homogeneous groups of economies and highlighting what differentiates them (Archibugi et al., 2009). R&D expenditure also not always result in the development of new products or processes, so it cannot be implied that the biggest spendings will result in the best innovative performance. Therefore, the next section of analysis will focus on findings from Global Innovation Index for four different groups of nations, divided based on their income levels.

	Global Innovation Index	Innovation Input Sub-index	Innovation Output Sub-index	Innovation Efficiency Ratio
High-income economies (47 in total)				
1	**Switzerland (1)**	Singapore (1)	**Switzerland (1)**	**Switzerland (1)**
2	**Netherlands (2)**	**Switzerland (2)**	**Netherlands (2)**	Luxembourg (2)
3	**Sweden (3)**	**Sweden (3)**	**Sweden (3)**	**Netherlands (4)**
4	United Kingdom (4)	United Kingdom (4)	Luxembourg (4)	Malta (7)
5	Singapore (5)	Finland (5)	Germany (5)	Hungary (8)
6	United States of America (6)	United States of America (6)	United Kingdom (6)	Germany (9)
7	Finland (7)	Denmark (7)	United States of America (7)	**Sweden (10)**
8	Denmark (8)	Hong Kong (China) (8)	Finland (8)	Estonia (12)
9	Germany (9)	**Netherlands (9)**	Ireland (9)	Ireland (13)
10	Ireland (10)	Canada (10)	Israel (11)	Israel (14)
Upper-middle-income economies (34 in total)				
1	**China (17)**	**China (27)**	**China (10)**	**China (3)**
2	**Malaysia (35)**	**Malaysia (34)**	**Bulgaria (34)**	Iran, Islamic Rep. (11)
3	**Bulgaria (37)**	**Croatia (42)**	**Malaysia (39)**	**Bulgaria (19)**
4	**Croatia (41)**	Russian Federation (43)	**Croatia (42)**	Turkey (25)
5	**Thailand (44)**	**Bulgaria (44)**	Turkey (43)	**Thailand (33)**
6	Russian Federation (46)	South Africa (48)	**Thailand (45)**	**Croatia (37)**
7	**Romania (49)**	**Romania (49)**	Iran, Islamic Rep. (46)	Costa Rica (43)
8	Turkey (50)	Colombia (50)	**Romania (48)**	**Romania (47)**
9	**Montenegro (52)**	**Montenegro (51)**	Costa Rica (51)	**Malaysia (48)**
10	Costa Rica (54)	**Thailand (52)**	**Montenegro (55)**	**Montenegro (56)**

	Global Innovation Index	Innovation Input Sub-index	Innovation Output Sub-index	Innovation Efficiency Ratio
Lower-middle-income economies (30 in total)				
1	**Ukraine (43)**	Georgia (53)	**Ukraine (35)**	**Ukraine (5)**
2	**Viet Nam (45)**	**India (63)**	**Moldova, Rep. (37)**	**Moldova, Rep. (6)**
3	**Moldova, Rep. (48)**	**Viet Nam (65)**	**Viet Nam (41)**	Armenia (15)
4	**Mongolia (53)**	**Mongolia (66)**	**Mongolia (47)**	**Viet Nam (16)**
5	**India (57)**	**Ukraine (75)**	Armenia (50)	**Mongolia (30)**
6	Georgia (59)	Tunisia (77)	**India (57)**	Kenya (41)
7	Tunisia (66)	**Moldova, Rep. (79)**	Georgia (62)	Egypt (45)
8	Armenia (68)	Philippines (82)	Tunisia (63)	Pakistan (46)
9	Philippines (73)	Morocco (84)	Kenya (64)	**India (49)**
10	Morocco (76)	Kyrgyzstan (85)	Jordan (67)	Jordan (50)
Low-income economies (15 in total)				
1	**Tanzania, United Rep. (92)**	Rwanda (73)	**Tanzania, United Rep. (71)**	**Tanzania, United Rep. (31)**
2	Rwanda (99)	**Uganda (98)**	Madagascar (85)	Madagascar (40)
3	**Senegal (100)**	**Nepal (101)**	**Senegal (90)**	Zimbabwe (69)
4	**Uganda (103)**	**Senegal (102)**	Zimbabwe (99)	**Senegal (70)**
5	Madagascar (106)	**Tanzania, United Rep. (106)**	Mali (100)	Mali (73)
6	**Nepal (108)**	Benin (110)	**Malawi (108)**	**Mozambique (88)**
7	Mali (112)	**Malawi (111)**	**Mozambique (109)**	**Malawi (89)**
8	Zimbabwe (113)	**Mozambique (112)**	**Uganda (111)**	Guinea (102)
9	**Malawi (114)**	Niger (113)	**Nepal (114)**	**Nepal (107)**
10	**Mozambique (115)**	Burkina Faso (117)	Guinea (118)	**Uganda (108)**

Figure 4: Countries ranked based on their income group (Global Innovation Index, 2018)

This ranking presents countries based on their performance in both the Global Innovation Index scores and the respective sub-indices. All 126 countries are divided into four income groups: high, upper-middle, lower-middle, and low. It is worth highlighting that the most (47) economies are classified as high-income and the least (15) belong to low-income group. In parentheses next to each country, there are the overall places in the ranking for each indicator.

It is said that the gaps between high-income economies, leading the innovation landscape, and less-developed countries remain big on nearly all innovation input and output metrics (Global Innovation Index, 2018). Both the overall score and innovation input sub-index include countries only from high-income economies in the top 10, while innovation output sub-index involves China on the 10th place, as the only representative of upper-middle-income economies. When it comes to the innovation efficiency ratio the ranks are more diverse. China ranks on the 3rd place in this measure. Additionally, it is the only indicator where countries from lower-middle-income economies are also present in top 10, as it can be seen with examples of Ukraine and Moldova, ranked on the 5th and 6th places respectively.

It can be clearly seen from the Global Innovation Index report that despite significant investments in innovation inputs, some economies do not generate a corresponding level of innovation outputs. There are important outliers that strongly over or underachieve in terms of innovation efficiency. Among high-income countries: Switzerland, Netherlands, Sweden, Luxembourg, Germany, Ireland, and Hungary stand out for producing significant innovation outputs for their given level of inputs. On the other hand, Australia, Singapore, Japan, Hong Kong, New Zealand, Canada, and Norway, as well as, many economies rich in resources, such as: Saudi Arabia, Qatar, or Trinidad and Tobago stand out as underperformers. Among upper-middle-income countries, China strongly overperforms in innovation efficiency. Lastly, among lower-middle economies, Ukraine, Moldova, and Viet Nam stand out as performing better than would be expected from their innovation inputs.

Over the last century, there has been a discussion and a certain focus on the European paradox, which referrs to the assumed fact that Europe has excellent higher education systems, good research infrastructure, and scientific research results but struggles to translate these assets into innovation and economic growth. Therefore, it has good performance in innovation inputs but worse in outputs. Similar statements have been argued to exist in countries such as: the Netherlands, Finland, and Sweden. But the studies have shown that some countries that performed well in terms of STI indicators did not perform sufficiently in terms of innovation, therefore it indicates that what has been registered as a paradox, might be more of a systematic weakness in the theoretical analysis and the indicators on which the analysis is built (Lundvall, 2017)(Global Innovation Index, 2018).

Actually, it can be seen from the table presenting the ranking for the high-income economies that the aforementioned countries are all among the top 10 innovative economies, according to Global Innovation Index. However, it is interesting to look more into the sub-indices for them. Starting with the Netherlands, which is placed on the second place in the overall ranking, it can be seen that the country scores high in both innovation outputs and innovation efficiency ratio, but performs worse in innovation inputs. Taking into account Finland, its overall rank is on the seventh place, taking 5th position in the innovation inputs, and 8th place in the innovation outputs sub-index. The most interesting figure about Finland is that the country is not included at all in the top 10 nations based on the efficiency ratio, having the same value of the indicator as the United States, equal to 0.76. Sweden, as the 3rd country mentioned in the studies about the European paradox, is in a different situation in the ranking than the previous two countries. Namely, Sweden is ranked on the 3rd place in both the overall index, and the two sub-indices for innovation inputs and outputs. Sweden's position in the efficiency ratio is 10th place, with a value of 0.82.

The next sections will look more into quantitative data, used to construct Global Innovation Index, followed by the use of different unsupervised learning methods to discover patterns and similarities between observations.

6.2 Exploratory Data Analysis

Exploratory Data Analysis (EDA) refers to the necessary process of performing initial investigations on the dataset to discover patterns, spot anomalies, test hypotheses, and to check assumptions using summary statistics and graphical representation of data. Exploratory Data Analysis is all about making sense of the data, before actually working on it. Understanding the data first and trying to gather as many insights from it as possible is considered to be a good practice (Patil, 2018).

As a part of Exploratory Data Analysis, the dataset was checked for missing values, as it is considered to be an important part of it. Missing values can affect the quality of machine learning models and sometimes need to be cleaned or removed before using predictive models (Laufer, 2019). From the way the index is constructed, as a consequence of missing values in some variables, the whole value range decreases for them, as the ranks for respective economies are stored as values in each column.

key <chr>	num.missing <int>
Entertainment and media market	63
Utility model applications by origin	61
Assessment in reading, mathematics, and science	54
Microfinance institutions' gross loan portfolio	46
Venture capital deals	46
Research talent in business enterprise	42
Cultural and creative services exports	40
Market capitalization	39
Firms offering formal training	35
GERD performed by business enterprise	35

1-10 of 63 rows   Previous 1 2 3 4 5 6 7 Next

Figure 5: Missing values from the dataset

The table above presents ten variables with the most missing values. The report includes a note that missing values were indicated with “n/a” and were not considered in calculating scores for sub-pillars, therefore they did not affect the scores negatively (Global Innovation Index, 2018). As it can be seen from the table, variable with the most missing values is representing entertainment and media market, with missing values for 63 countries which equals to 50% of all observations. The main reason for that might be that this variable was introduced in Global Innovation Index only in 2014, as an attempt to include broader sectoral coverage under the creative outputs of an economy. The variable with second highest value of missing observations represents utility model applications by origin, which is also a part of innovation outputs. Assessment in reading, mathematics, and science, with 54 missing values, is a part of human capital and research sub-pillar, under innovation inputs. When the following numbers of missing values are considered, it can be seen that most of them are for variables under business and market sophistication sub-pillars. To obtain more information about a percentage of missing values for all variables that contain such, Figure 13 is included in the appendix.

The next part of Exploratory Data Analysis will be looking into correlations between chosen variables. The plot will show correlation coefficients between the main ranking, innovation inputs, outputs, as well as, the seven main pillars of Global Innovation Index.

It is important to note that the main pillars refer mostly to technology and capacity competitiveness, described in the theoretical part of the project, which relates to abilities to compete successfully in different markets with new goods and services, and capabilities that are important for exploiting technological opportunities. Technological competitiveness is measured by indicators related to R&D, patents, and ICT infrastructure, while capacity competitiveness is reflected in variables related to education, availability, and quality of financial institutions or markets, and quality or efficiency of governance (Fagerberg et al., 2007).

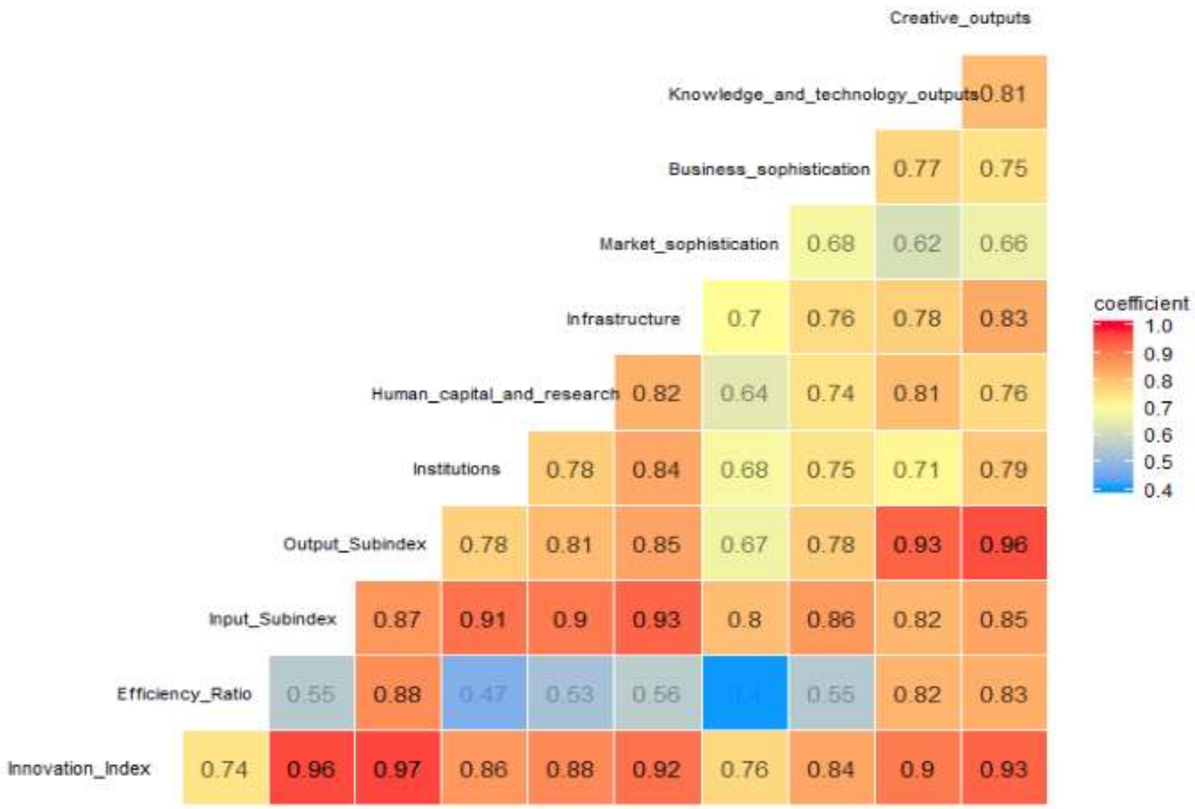

Figure 6: Correlation matrix

From the way the framework is constructed, it was obvious that all the correlations will be positive, as all the values in the dataset are positive and the value range for most of the variables is 1 to 126. Thus, the colors are also scaled to reflect the specific data and the lowest values (0.4) are associated with blue, while the highest with red, as usually.

From the visualization, it becomes clear that the strongest correlation is between the main ranking and innovation outputs, followed by innovation inputs. More interesting is the fact that the correlation coefficient between positions in the Global Innovation Index and Innovation Efficiency Ratio is equal to 0.74, from which it can be concluded that the performance of countries is determined more by how much inputs or outputs does the economy create towards being innovative, rather than measuring how efficiently the inputs are being used. Therefore, it is possible that some nations will be highly ranked based mostly on a good environment for innovation that they have or high investments in innovation, not necessarily using them efficiently. There are some potential drawbacks of this approach to innovation measurement and constructing composite indicators in this way. Most importantly, some decision-makers can put emphasis on producing a significant amount of innovation inputs, just to increase the value of this subindex, which would make the country's performance much better, without actually improving the situation within or creating any innovations.

The efficiency ratio itself has stronger positive correlation with innovation outputs than inputs, which means that in order to be ranked highly in this index, countries should focus on producing innovations rather than just investing in them. Thus, the measure reflects the actual measured outcomes of the innovation activities.

In general, the highest coefficient including the main pillars occurs for innovation outputs and creative outputs (0.96). The efficiency ratio shows the weakest correlations with most chosen variables, especially if it comes to innovation inputs, which implies that changes in the environment for innovation or new investments in innovation activities, do not have a strong effect on innovation efficiency. Situation on the market and institutional set-up seem to affect efficiency ratio the least, being also the lowest values in the whole matrix, equal to 0.4 and 0.47 respectively.

Both the main ranking and innovation inputs sub-index are strongly correlated to infrastructure. This might lead to two conclusions. Firstly, it could be assumed that the countries which have the best infrastructure might also be the most prone to generating innovations. Or secondly, looking from a different perspective and noticing potential drawbacks, there is a risk that decision-makers in some regions could influence the rankings by focusing on the areas that contribute to better innovative performance stronger, and omit some other areas that need improvements. There could also be a risk of some bias in the overall ranking. If, for instance, some country has a very good infrastructure, it will be ranked higher but in reality, it does not necessarily imply being more innovative. Good environment for innovation does not imply that firms will use it efficiently and produce a lot of innovations.

The sub-pillar, which is the least correlated with any other chosen variables, is market sophistication with the value range from 0.62 - 0.8. It suggests that the changes within this group of variables affect the overall score in GII the least, even though it was said that access to well-developed capital markets is essential for developing innovative concepts (Fagerberg, 2013).

The subset of data included in the second correlation matrix, which includes only the scores for each index and the main pillars building up the ranking, will be also used for the next section of data analysis - dimensionality reduction and clustering. The only exception will be that the main ranking stored in the variable Global_Innovation_Index will be left out.

6.3 Dimensionality reduction

Principal Component Analysis (PCA) is a part of exploratory data analysis. This method allows visualizing the variation present in a dataset with many variables, and helps to see the overall "shape" of the data, identifying which samples are more similar to each other, and which ones are different (Hayden, 2018). The idea behind it is that strongly correlated variables refer to the same underlying dimension, so that a dataset containing many variables can be reduced into a small number of principal components, which reflect a significant part of the total variance (Fagerberg et al., 2007).

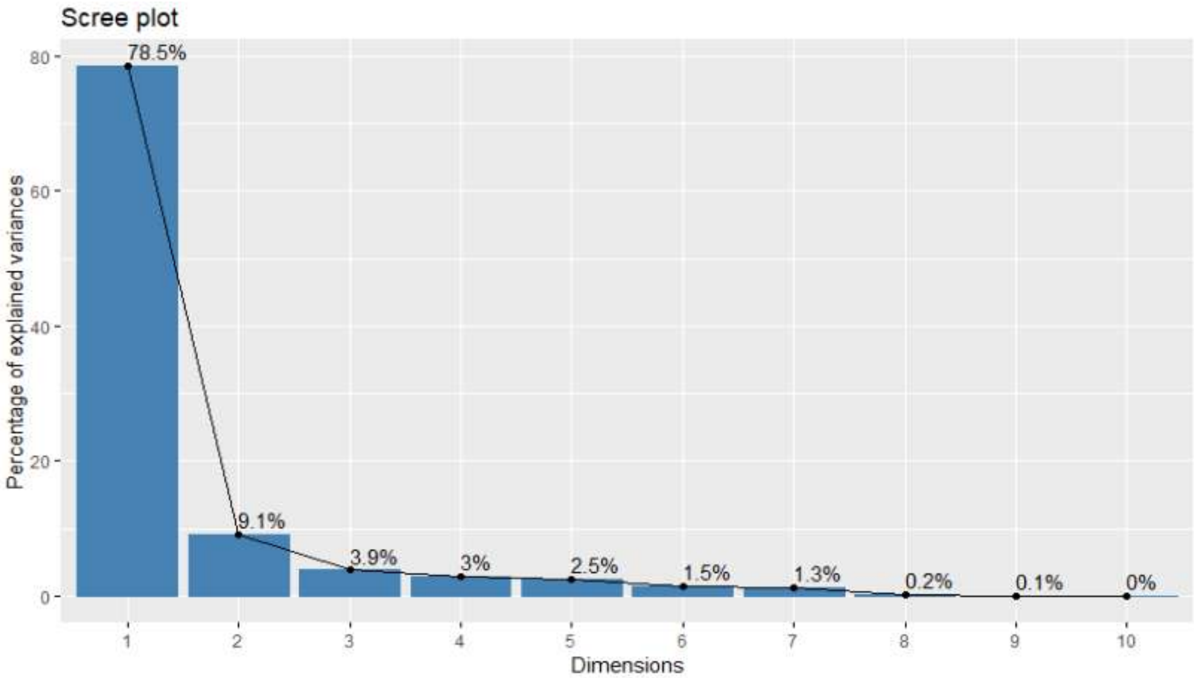

Figure 7: Scree plot for chosen variables

In case where many variables correlate with each other, they contribute strongly to the same principal component. Each principal component sums up a certain percentage of the total variation in the data. The percentages shown for each dimension can be accumulated, which represents the marginal effect of adding respective components. Thus, there is a trade-off between explaining more variance in the dataset and having more dimensions. In general, the number of principal components equals to the number of variables that are included in the whole dataset, but if the last 10 components explain only small percentage of variance cumulatively, then including them would not add much value to the analysis. The aim should always be to reduce dimensionality and set for the minimal number of components, which would explain satisfactory proportion of the variance. If the initial variables are strongly correlated with each other, it would be possible to approximate most of the complexity in the dataset with just a few principal components (Hayden, 2018).

As it could already be seen from correlation matrices for Global Innovation Index dataset, most variables have strong positive correlation with one another. Therefore, a big proportion of the variance is already explained within the first principal component. The percentage significantly decreases from the second component. In the case of including only the main sub-pillars of GII, together with all the indices, based on which the ranking is constructed, the first component explains 78.5% of the total variance already. When the second component is added, the value increases to 88.6%. As it is a satisfactory threshold already, further analysis will be conducted on the two first principal components. The same steps were repeated for the whole dataset, including all indicators, but the values were only 46.3% for the first component and 52.2% for the first two components combined (see Figure 14 in Appendix).

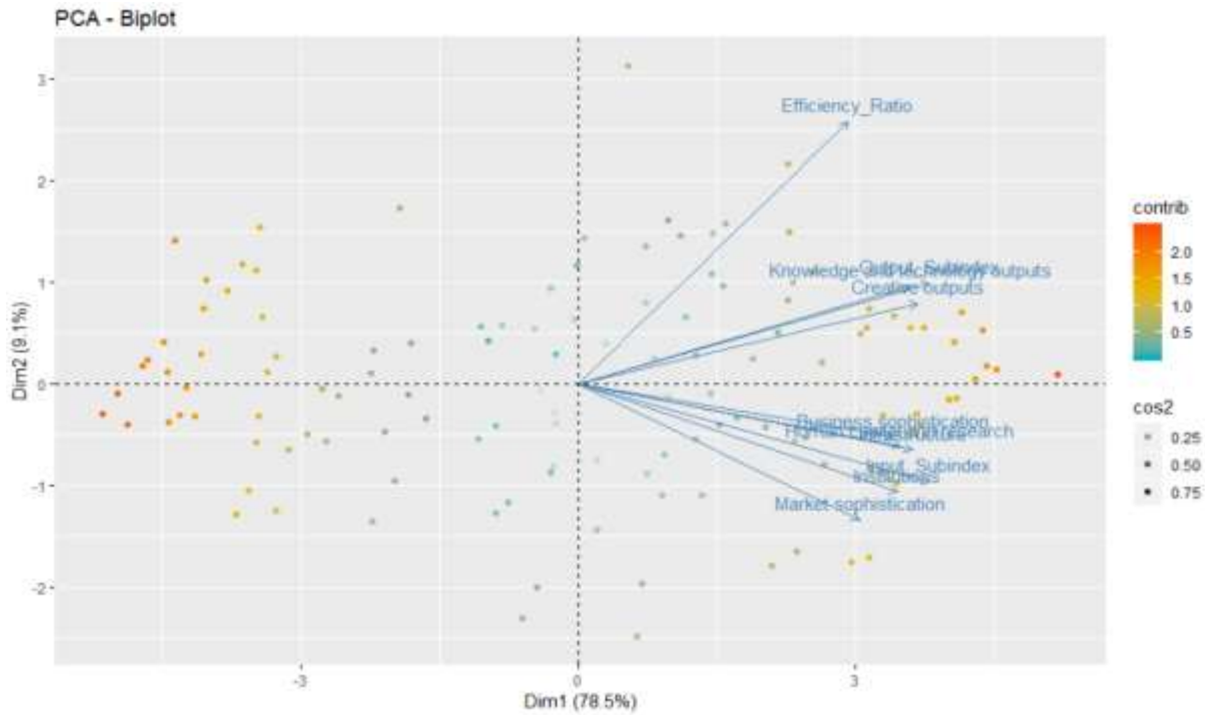

Figure 8: Biplot for chosen variables

A biplot is a type of plot allows visualizing how the samples relate to one another - which samples are similar and which are different, and simultaneously reveals how each variable contributes to each principal component. It is said that in order to interpret each principal components, it is needed to examine the magnitude and direction of the coefficients for the original variables. The larger the absolute value of the coefficient, the more important the corresponding variable is in calculating the component (Hayden, 2018). The direction to which variables are pointing shows whether the contribute positively or negatively to each principal component. Another characteristic of biplot can be summarized as the closer variables are together on the plot, the more similar they are. So, it is also possible to identify clusters of similar indicators from the biplot.

This biplot shows both the variables, together with their contribution, and observations (countries) represented by dots on the plot. Another biplot containing only the variables can be seen in the Appendix as Figure 15. As can be seen from the plot, all analyzed variables contribute positively to the first component. They are more differentiated if it comes to the second component. In this case, it is possible to distinguish two groups. Variables that contribute positively to the second principal component include efficiency ratio, output subindex, knowledge outputs, and creative outputs. They are all related to innovation outputs, and as it is known already from analysis of correlations, efficiency ratio also depends on the outputs, therefore it is clustered in this group. This is also an indicator that has the strongest positive contribution to this component, with market sophistication on the other side of the spectrum with the strongest negative contribution. From the analysis of correlations it was already clear that those two variables had the lowest correlation coefficient in the whole matrix, so it should be expected that they will be plotted at the opposite sides. The second group of variables, contributing negatively to the second principal component, includes all variables related to innovation inputs.

Observations are spread out throughout the whole plot, where the ones in shades of dark orange represent countries that score the highest or the lowest within chosen criteria. There are also a few outliers, showing that some countries outperform the others in, for example, efficiency ratio and creation of innovation outputs, or conversely, some nations may have exceptional innovation inputs.

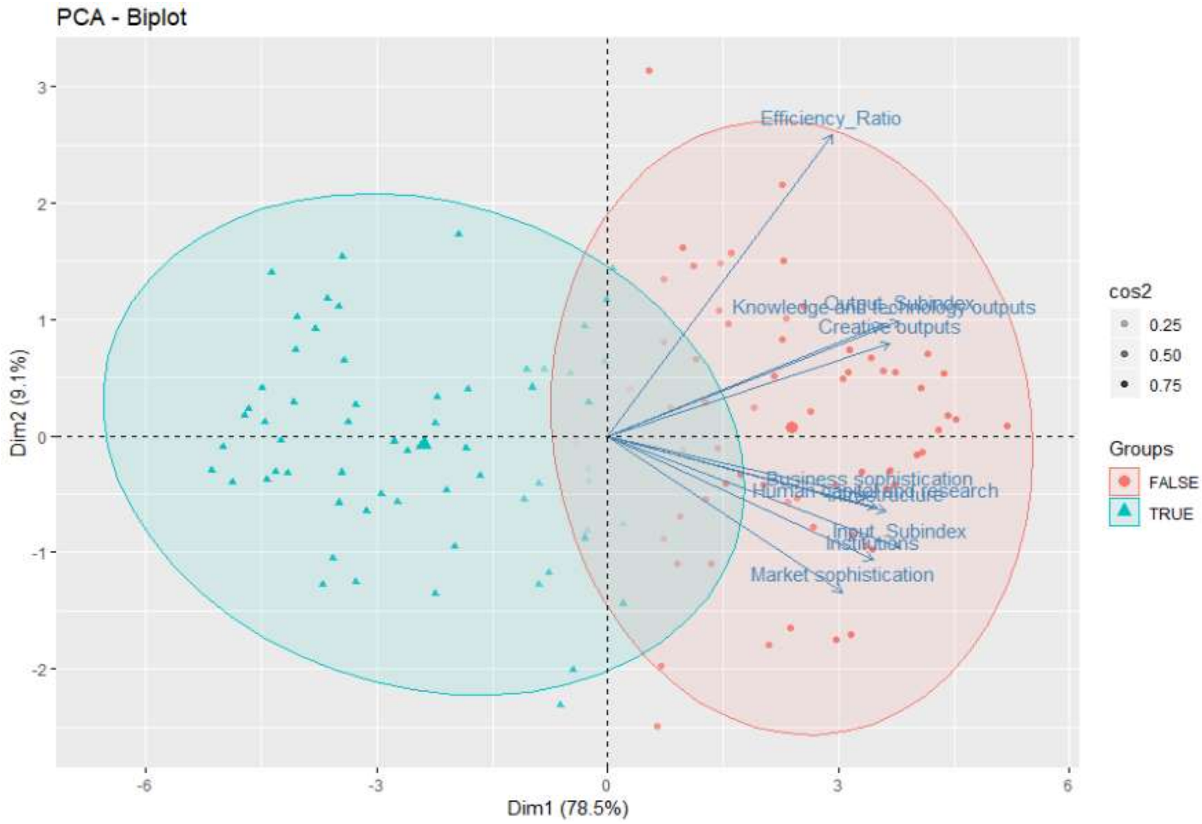

Figure 9: Biplot divided by mean scores in the ranking

From all biplots combined together, it is almost certain that the first principal component is related to innovativeness, or the overall ranking itself. Therefore, all the sub-indices and sub-pillars of Global Innovation Index contribute positively to this dimension.

The last biplot represents the same variables as the previous visualization, with one exception. This time two groups of observations were plotted based on the mean value of Global Innovation Index. They have been differentiated in a way that the blue group on the left represents observations that have ranks lower than the mean value of Global Innovation Index, and the red group represents remaining observations. However, it is worth noting that the main index represents ranks, thus the bigger the value, the worse the performance of an observation, which means that economies placed on the right side of the plot are considered less innovative.

From the ways observations are distributed it becomes clear that some score well in either innovation inputs or outputs, which places them in one of the groups above but in order to gain some understanding from this, it would be more relevant to look at specific countries. Consequently, the next visualization will present countries plotted based on the same criteria on the biplot below.

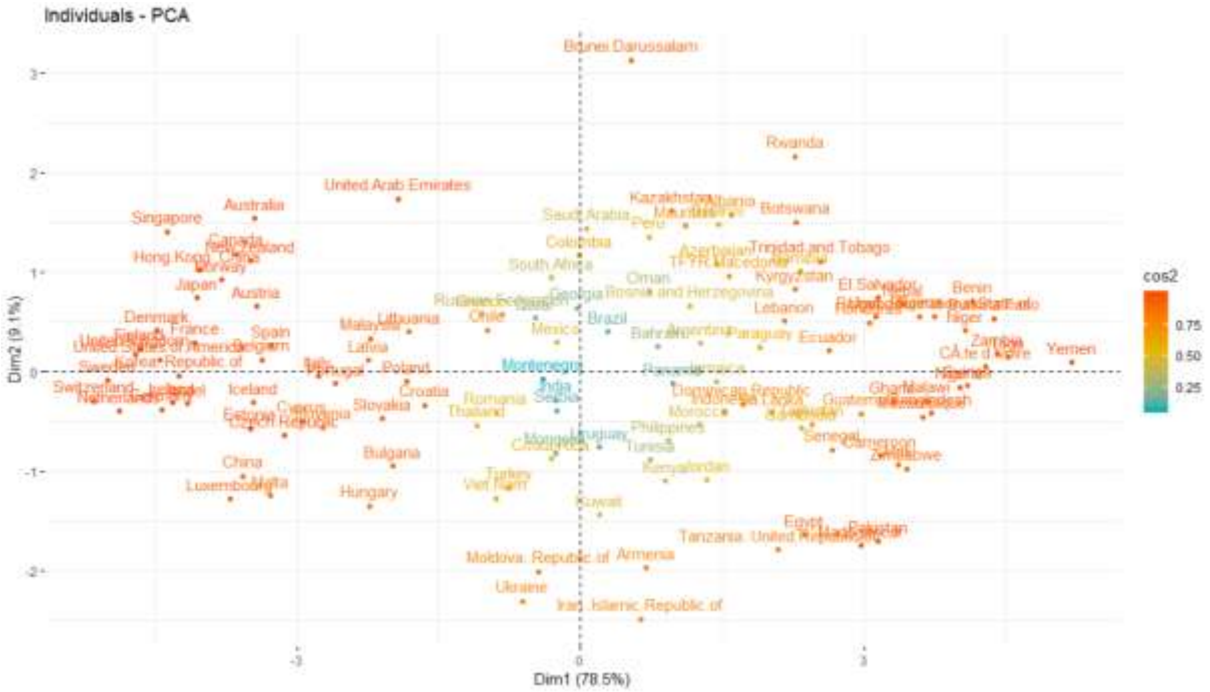

Figure 10: PCA for individual countries

Now it is possible to analyze characteristics of individual economies, which would lead to further understanding why each observation was placed at a certain point of the plot. It would be the most interesting to check features of economies that are different, for example some extreme values, outliers, or the most central points.

The first outlier which is most visible from the plot is Brunei Darussalam. It is placed on the 67th rank in the overall ranking, which would not make it particularly interesting case to analyze. What is more interesting about the country, is that is scores extremely low in the innovation efficiency ratio, being placed on the 124th position, with 37th rank in innovation inputs, and 112th place within innovation outputs. It scores quite well within the indicators under institutions pillar, being placed on the 6th rank for the political stability, and on the first place for the cost of redundancy dismissal. The low cost of redundancy is good for the businesses, but on the other hand, if it is high, it means that the employees are more protected. Brunei has also quite good positions within the human capital & research pillar, with 14th place for pupil/teacher ratio, and 8th place for the graduates in science & engineering. However, the score for tertiary enrollment in general is not impressive with 79th rank, thus a big percentage of engineers among not so many students, might not make much of a difference. Brunei Darussalam is placed on 18th rank for the general infrastructure, and 9th for gross capital formation, 17th for the market sophistication pillar, with 2nd place in ease of getting credit. Therefore, it seems like the country has a good environment for innovation and starting businesses. Nonetheless, it scores low on the innovation outputs, except for sub-pillar related to knowledge workers, where Brunei takes 25th place. The country has 124th place for knowledge absorption, 113th for knowledge and technology outputs, and 105th for creative outputs. There is quite a lot of missing data about creative outputs, culture, and entertainment for the country, which could also affect the scores.

The next country that will be analyzed is Montenegro, as the most central point of the visualization. It is placed on the 52nd position in the overall ranking and from the assumptions made about the dimensions of the two principal components, it should have rather central scores for most pillars. The country's efficiency ratio is on the 56th place, innovation index on the 51st, and innovation outputs on the 55th. Mostly, indeed the scores are close to the median, with a few exceptions. Montenegro is placed on the first rank for country-code top-level domains, and on the 4th for printing, publications & other media output, both parts of creative outputs. It also scores high in the ease of getting credit (11th place), ICT services imports (12th place), and foreign direct investment net inflows on the 10th rank. It is interesting that the economy got 119th place in the FDI net outflows, which means that there more external economies are investing in Montenegro, than the country invests abroad. There are also some indicators that the country scored exceptionally low in, such as 121st place for knowledge diffusion, 123rd rank for trade, competition & market scale sub-pillar, with the last place for domestic market scale, which is supposed to address the impact that the size of a country has on its capacity to introduce and test innovations on the market. Montenegro does not perform badly in other indicators under the market sophistication pillar. The main pillar, which is considered an advantage for the country is related to institutions (46th overall place), with main strength in business environment on the 35th position. Montenegro also has good score in tertiary education, with the 23rd place and 56.9% of population enrolled in universities. There is missing data for graduates in science and engineering, and students mobility, which could affect the scores for the whole sub-pillar. Montenegro also scores high (28th place) in scientific and technical articles, which are a part of knowledge & technology outputs. It also has good performance in indicators related to information & communication technologies, with 39th place for the general ICTs infrastructure, and 22nd place for computer software spending.

The next country worth looking more into is Rwanda (upper right corner of the plot). It is positioned on the 99th place in the overall ranking. If it comes to different sub-indices, Rwanda is on the 73rd rank in innovation inputs and 120 in outputs, resulting in 125th position for innovation efficiency ratio. Rwanda scores low in the human capital & research pillar, with the 107th position. It consists of 81st place in education, with 92nd rank for expenditure on education as a percent of GDP, and school life expectancy equal to only 11.2 years. The overall sub-pillar for tertiary education places Rwanda on 116th position, with 108th place for tertiary enrolment (only 8% of population). However, Rwanda scores well for government's funding per secondary student on the 5th rank overall, which might improve the scores related to education in the future. Low scores in inputs related to education, directly affect innovation outputs. Knowledge-intensive employment is ranked as 114th, there is also a very low value for female employees with advanced degrees (0.7%), placing Rwanda as 102nd. Nonetheless, 55.4% of firms offer formal training, placing Rwanda as 11th in the overall ranking. The reason for that could be that tertiary education is not so popular and firms train the workforce in lacking skills. The sub-pillar related to innovation linkages has also good score (36th place), consisting of University/industry research collaboration (54th place), and clusters development (41), which could indicate that even though there are not so many educated people, the country is trying to take advantage of knowledge sharing and use available resources.

The overall score for R&D sub-pillar is also low, with 117 place, however the indicator measuring global R&D expenditure by top 3 companies places Rwanda as 40th. This means that even though public expenditure on research and development could be improved, businesses still aim at being more innovative. The country is placed on 91st rank for infrastructure pillar, with the main strength in for general infrastructure (40th place), and main weakness related to Information and communication technologies, with overall 104th rank, 116th place for ICT access, and 111th for ICT use. This is directly related to low scores in online creativity, as a part of innovation outputs, where Rwanda is placed as 122nd, as well as, 124th place for knowledge and technology outputs that contain computer software spending, and high-tech net export. However, because ICTs sector is not so developed, indicator related to high-tech net imports was identified as economy's strength, placing Rwanda on 40th rank, and shifting knowledge absorption sub-pillar to 72nd place. Considering other advantages, Rwanda is ranked 6th for the ease of getting credit, first for gross loan portfolio of microfinance institutions, and 24th for the whole investment pillar, meaning that it is a good place to attract capital and foster entrepreneurship.

Another country that could be an outlier is United Arab Emirates. It has 38th place in the overall ranking, with 24th rank for innovation inputs, 54th for innovation outputs, giving the country 95th place for efficiency ratio. Even though, it belongs to high-income economies and has very good infrastructure (12th place for general infrastructure sub-pillar), the country does not produce corresponding level of innovation outputs. UAE is ranked 15th for education, and 9th for the tertiary education but it does not translate to direct outputs related to knowledge creation. The country is placed on 100th position for scientific & technical articles, despite there is 22%of graduates in science and technology. However, it can be explained by another indicator - research talent as a percentage in business enterprises, which is equal to 62.2% and places United Arab Emirates on the 8th rank. Thus, local researchers probably do not write so many articles but rather work in the private sector. It can also be seen that the country is not characterized by any protectionism, as all the indicators related to patents, trademarks, and industrial designs are quite low. Therefore, it might be an explanation why UAE scores second place in the state cluster development, as the technologies are not protected, fostering knowledge diffusion.

The next country to be analyzed is Singapore, situated in the upper left part of the plot. The overall rank of Singapore is 5th, with first place for innovation inputs, and 15th for innovation outputs, resulting in 63th rank for the efficiency ratio, indicating that the country is underperforming in producing innovations. Singapore has the first place in institutions pillar, winning in all indicators related to political environment, and regulatory environment, except for 9th place for the rule of law. This indicator is supposed to reflect to which extent agents have confidence in the rules of society, and in particular the quality of contract enforcement, property rights, police, and courts, as well as, the likelihood of crime and violence. Although the political scene in Singapore is dominated by the People's Action Party (PAP) since 1959, it is considered the least corrupt country in Asia and in the top ten in the whole world.

The scores for business environment sub-pillar are lower than other parts of institutional set-up, placing Singapore on the 19th rank, with 25th place for the ease of resolving insolvency, and 6th for the ease of starting a business. The country is also leading in human capital & research pillar, with first place in reading, maths & science assessment for pupils, and tertiary education sub-pillar. Unfortunately, there is missing data for both tertiary enrollment and percentage of graduates with science or engineering degrees, preventing more in-depth analysis. However, Singapore is placed on the 5th rank in number of researchers per million citizens, and second for the knowledge-intensive employment, indicating that there are many citizens with advanced degrees or that the country attracts foreign researchers. The economy scored 5th rank for the overall infrastructure with the main strengths laying in Government's online services (3rd rank), logistics performance (5th rank), E-participation (8th rank), and GDP/unit of energy use on the 9th place. The main disadvantages include 22nd place in ICT use, 36th in gross capital formation as the percent of GDP, 40th in ISO 14001 environmental certificates, and 45th in environmental performance. It is quite known that Singapore has issues with pollution. However, the country recently made heavy investments in renewable energy, especially photovoltaic solutions, taking advantage of research capacities within the country and open innovation. These steps might result in both improving the environmental indicators, and achieving more patents and inventions related to clean energy.

Singapore is ranked on the 4th place for market sophistication, leading the overall ranking in indicators concerning market capitalization, and applied tariff rate, as well as, scoring the second place for investments sub-pillar. The main weaknesses within this pillar are 26th rank for the ease of getting credit, and 38th for domestic market scale. Singapore is situated on the 53rd place for foreign financing as a percentage of total expenditure on R&D. Within innovation linkages sub-pillar, it is ranked 3rd for joint ventures/strategic alliance deals, and 8th University/industry research collaboration, fostering knowledge exchange. Singapore is ranked on the 2nd place for overall knowledge absorption sub-pillar, with 1st rank for both charges for use of intellectual property as a percent of total trade, and for FDI net inflows. It is interesting to note that Singapore leads in foreign direct investments, while the gross expenditure on R&D from abroad is quite low, compared to other economies.

If it comes to innovation outputs for Singapore, it is placed on the 11th position for knowledge & technology outputs, and on the 35th for creative outputs. The main strengths are first places for: high & medium-high-tech manufactures, high-tech net exports, and FDI net outflows. Interesting to see that the country leads both in inflows and outflows of foreign direct investments, meaning that there is both favorable environment that attracts foreign capital, and that Singapore seeks other markets to develop some technologies or products. Even though the overall score for the economy for the knowledge diffusion sub-pillar is quite high (4th rank), the indicator it consists of, measuring ICT services exports, is a significant weakness, placing Singapore on the 69th position. Other low scores include 88th place for the number of trademark applications, 62nd for industrial designs, 70th in printing & other media as a percentage of manufacturing, and 38th for national films. Generally, the country scores low for the creative outputs and within knowledge creation sub-pillar. Creativity can be explained by the fact that other industries and sectors are promoted by the government, with more focus on high-tech, as the best scoring indicators within this pillar are related to ICTs & business model creation, and ICTs & organizational model creation. Both indicators are measured by answers to surveys, based on questions about to what extent do ICTs enable new business models or new organizational models. When knowledge creation is considered, the lowest score is associated with patents applications. This might be a positive characteristic, associated with low protectionism, that can be a factor which attracts foreign investments in the country. Surprisingly, Singapore is also ranked 27th for the scientific & technical articles, which is quite a low score taking into consideration a high number of researchers in the country.

Another economy worth examining is Islamic Republic of Iran, situated as the lowest point on the plot. Iran is ranked on the 65th position in Global Innovation Index. Next country, located close to Iran in the lower part of the plot, is Ukraine. Ukraine scored 43rd place in the overall ranking, being more than 20 ranks ahead of Iran, thus it is worth checking whether the two economies are similar to each other based on their position on the plot, and if yes, in which matters. Ukraine is an example of great performance in terms of innovation efficiency ratio, being placed on the 5th position, with 35th rank for innovation outputs, and 75th in innovation inputs. The country moved up 7 ranks compared to the previous year. The biggest similarity between countries stands within producing high level of innovation outputs, given their level of inputs. Iran is on the 11th place for innovation efficiency, 46th for innovation outputs, and 93rd for inputs. Both countries also belong to upper-middle income economies. What is also interesting, both countries score similarly in the main pillars, being characterized by different strengths and weaknesses among particular indicators. For instance, Ukraine is placed as 107th in institutions pillar, while Iran is 110th. Similarly, for human capital & research, Ukraine is ranked 43rd, while Iran 45th, and for infrastructure the countries rank 89th and 87th respectively. The main differences are within market sophistication (Ukraine on 89th place and Iran on 106th), and business sophistication, where Ukraine ranks 46th, and Iran 108th. Those pillars could be used to explain the gap between countries.

Another country to be presented in this section is Hungary, situated in the lower-left corner. It is ranked 33rd in GII, 25th for innovation outputs, and 41st for innovation inputs, giving the country 8th place in efficiency ratio. Hungary is one of the countries that perform better than expected in innovation outcomes. The country is ranked 40th for the institutions pillar, with scores ranging from 32nd for political stability, to 65th for the ease of starting a business. It is on the 38th rank for human capital & research pillar, with the main strengths laying in pupil-teacher ratio (29th place), tertiary inbound mobility (23rd place with 8.9% of foreign students), and 34th place for the whole R&D sub-pillar. The country performs worse in indicators related to infrastructure. The 49th place for the whole pillar is based on good score (28th rank) for ecological sustainability, 60th in information & communication technologies, and 65th for general infrastructure. The main strengths lay within ISO 14001 environmental certificates (10th place), and 33rd rank for ICT access. While the main weaknesses are related to e-participation (89th rank), gross capital formation (93rd), and electricity outputs in kWh per capita (60th). It can be seen that even many citizens have access to information technologies, online tools are still not used to promote interactions between citizens and government. Also, even though the country performs well in measures related to the environment, the usage of electricity is still high.

From all input pillars, Hungary scores the lowest in market sophistication, with 86th place. The ranks for subsequent sub-pillars are as following: 90th for credit, 93rd for investment, and 54th for trade, competition & market scale. Surprisingly, it is relatively easy to receive credit (26th rank) but Hungary is 90th in indicator related to domestic credits to the private sector, and 92nd in ease of protecting minority investors, measuring transparency of related-party transactions, and shareholders rights in conflicts of interests. Additionally, the country ranks 117th in the intensity of local competition. Thus, taking into account all the risk, entrepreneurs might not want to start new businesses, even though the capital is easily available. Hungary scores exceptionally high in knowledge absorption sub-pillar (4th overall rank), affected by 5th place in FDI net inflows, 12th place in research talent as a percentage of employees in business enterprises, and 15th place in high-tech imports. It is also placed on the 22nd position for both gross expenditures on R&D financed and performed by business enterprises. This may indicate that the R&D is influenced by the private sector, or multinational companies that invest in Hungary, as 15% of global R&D expenditure is financed by abroad.

Hungary scores well in innovation outputs, given the economy's inputs. Its overall rank for knowledge & technology outputs is 16, with first place for FDI net outflows, 6th for high & medium-high-tech output as a percentage of total manufactures, and 9th for high-tech net exports as a percentage of total trade. The main weaknesses consist of ICT services exports (60th place) and growth rate of GDP per worker (87th place). It can be seen that many technological products are manufactured in the country but there is not so much focus on services. Additionally, Hungary invests a lot of capital abroad. If it comes to creative outputs, Hungary scores lower (44th place), with main strength in creative goods (8th place) and services (22nd) exports. It is worth noting that even though the country exports a lot of innovation outputs, the conditions for local employees are not improving, as the GDP per person employed has increased only by 0.2%. Therefore, Hungary might be the case that despite of producing many innovation outputs for given environment, they do not really benefit the country, as most of products are exported and the capital is invested abroad.

The last country, situated the most to the right of this plot and taking the last position in the whole ranking, is Yemen. It scores the lowest in both innovation inputs and outputs, however its efficiency ratio is situated on the 122nd place. The economy has also the last rank for institutions, business sophistication, and knowledge & technology outputs pillars, and 125th place for infrastructure. It scores particularly low in the indicators related to political environment, general infrastructure, and credit. However, there are some areas where Yemen performs better. For instance, it is ranked 5th for GDP per unit of energy use, which shifts the country to 41st position in ecological sustainability. Other favorable indicators include: average expenditure on R&D of the top three global companies (40th place), expenditure on education, which is equal to 5.2% of GDP, placing Yemen as 45th. Additionally, tertiary-level inbound mobility rate of 4.3% of all students being from abroad, situates Yemen on 46th place. If it comes to innovation outputs, the country performs the best (28th place) in charges for the use of intellectual property as a percentage of total trade, and has another strengths being 63rd in entertainment & media market, 67th in ICT services exports, and 66th in trademarks by origin. Otherwise, low scores in innovation inputs translate directly in the lack of innovation outputs.

After looking more in-depth into different characteristics of countries, it became clear that economies that scored higher in the overall ranking are situated more to the left on the PCA plot. Additionally, countries in the lower part of the plot have higher scores in innovation outputs, given their inputs, thus achieving higher efficiency ratio. Similarly, countries plotted in the upper part of the visualization can be considered underachievers.

6.4 Clustering

In order to analyze the data more in-depth, principal component analysis will be used to form groups of observations to detect some similarities or differences between countries. This section will be helpful to interpret some of the results, but above all, it can be used for better comparison of economies, and detecting common characteristics within groups of observations.

Clustering is a subjective task. Therefore, there can be more than one correct clustering algorithm for the specific problem, as every algorithm follows a different set of rules for defining similarity among data points. The clustering algorithm, which is most appropriate for a particular task often needs to be chosen experimentally, unless there is some mathematical reason to prefer one clustering algorithm over another. An algorithm might also work well on a specific dataset but fail for a different one (Jaiswal, 2018).

For this analysis, hierarchical clustering was chosen, as it gave better results than the simple centroid-based algorithms. In order to decide the rules for clustering, there are several ways of measuring the distance between clusters, called Linkage Methods. A few were tried to find the best way of performing this clustering (see Figures 19, 20, 21 in Appendix). The most common methods contain: complete-linkage (the maximum distance between clusters), single-linkage (the minimum distance between the clusters), average-linkage, and centroid-linkage, based on finding centroid of each cluster, and calculating the distance to it. It is said that different linkage methods lead to different clusters and there is no universal method (Pathak, 2018).

cluster <int>	Efficiency_Ratio <dbl>	Input_Subindex <dbl>	Output_Subindex <dbl>
1	73.25532	63.61702	69.48936
2	86.30769	106.79487	100.02564
3	29.80000	21.15000	20.85000

cluster <int>	Institutions <dbl>	Human capital and research <dbl>	Infrastructure <dbl>	Market sophistication <dbl>	Business sophistication <dbl>
1	66.21277	62.61702	65.19149	66.23404	73.89362
2	102.61538	104.46154	103.71795	93.94872	94.17949
3	22.17500	24.60000	22.30000	30.60000	21.37500

cluster <int>	Knowledge and technology outputs <dbl>	Creative outputs <dbl>	n <dbl>
1	70.34043	70.65957	47
2	97.87179	98.07692	39
3	21.95000	21.37500	40

Figure 11: Characteristics of clusters

Firstly, it is possible to examine what is inside of different clusters. The numbers show mean values for all chosen variables, grouped by a cluster. It becomes quite clear that the countries were grouped based on their performance in different pillars of Global Innovation Index. Cluster number 2 seems to have the highest ranks in all pillars, therefore it consists of countries from the lowest places in the ranking. As it can also be seen 39 economies belong to this group. Cluster number one is the biggest, containing 47 countries, for which the scores are close to the middle values for all chosen variables. And the remaining cluster number 3 consists of the economies that score the highest in all ranks. It can also be seen that the lowest mean value from the whole summary is for the output sub-index for the 3rd cluster, meanwhile the highest value is associated with input sub-index for the 2nd cluster. This means that countries clustered in the 2nd group scored exceptionally low in the innovation inputs, and those that belong to the 3rd cluster have very high ranks in innovation outputs. Additionally, the means are more spread out for some pillars. There is a clear distinction between the two sub-indices, as well as, the first three pillars. But innovation efficiency ratio and the remaining variables, on the other hand, are less distinguished between the clusters. While some of the indicators could be affected by missing values and countries could obtain higher scores just because there was missing data and the ranking went up to 80, instead of 126.

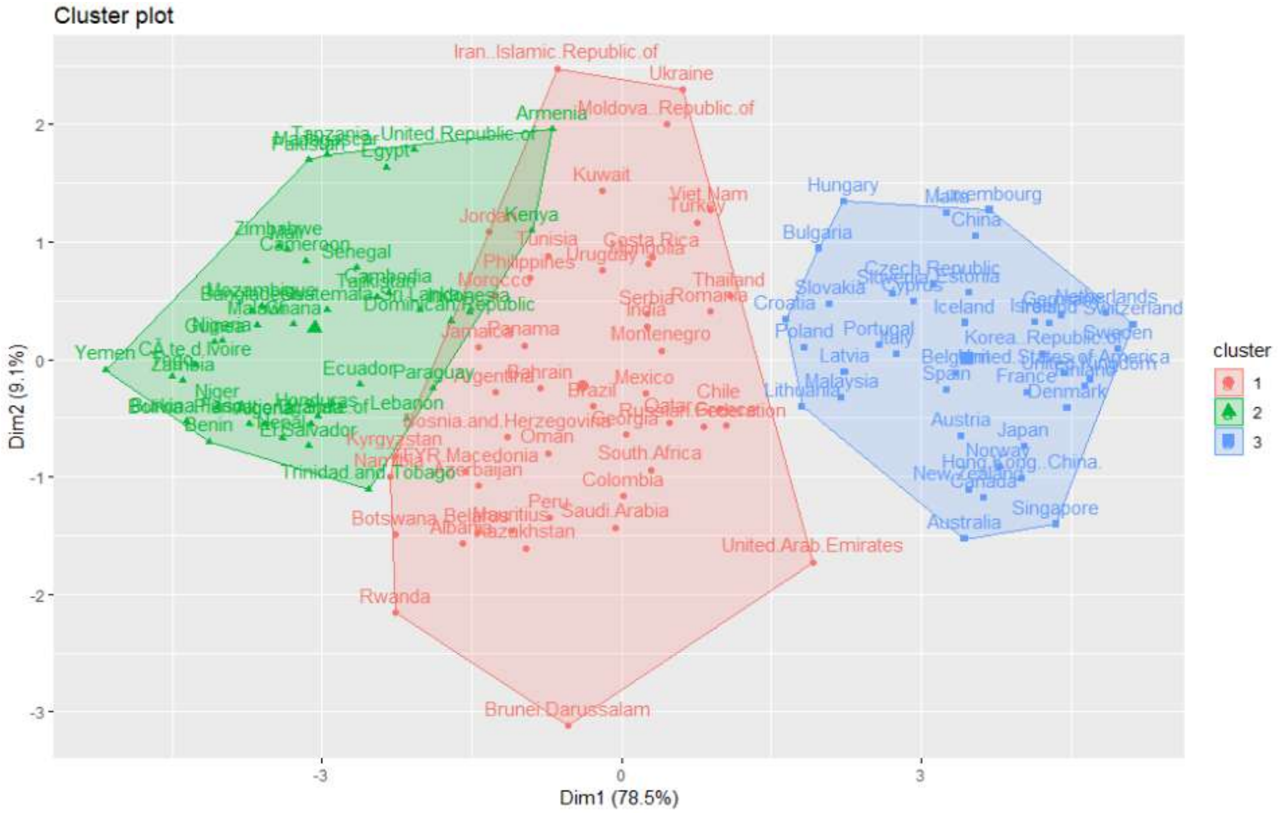

Figure 12: Hierarchical clustering plot

The similarity between clusters is often calculated from the dissimilarity measures. The larger the distance between two clusters, the better the clustering algorithm performs. Similarity also reflects the strength of a relationship between two data objects. In this case, the closer the observations are on the visualization, the more similar are their characteristics. There are many distance metrics used to calculate the dissimilarity measure, and their choice depends on the type of data. For instance, if it consists of continuous numerical values, the most popular distance metric, Euclidean distance, can be used (Jaiswal, 2018). The first characteristic of the visualization above is that there is one cluster completely separated from the other ones, and the two remaining clusters overlap. It can also be seen that the countries illustrated as outliers through PCA in the previous section are also the outermost points in clusters. In the following part, a few pairs of countries will be compared, in order to see what are the similarities and differences between them.

The first pair of countries will be Kenya and Jordan, which are situated very close to each other (upper middle part of the plot) but belong to different clusters. Both countries are plotted in the area, where two clusters overlap. Moreover, Kenya is more on the right in the two-dimensional space than Jordan, which might suggest that it could also belong to the other cluster. Considering the ranking in Global Innovation Index, the two economies are also very close to each other - Kenya is placed on the 78th rank, and Jordan on the following 79th. If it comes to sub-indices, Kenya ranks 91st for innovation inputs, 64th for outputs, and 41st for efficiency. Jordan scores 88th for innovation inputs, and 67th for innovation outputs, with efficiency ratio on 50th place. This suggests that the two countries are not so similar to each other, as Kenya outperforms Jordan for both innovation outputs and efficiency, while Jordan has better inputs. From the scores, it can be seen that Jordan is ranked higher for the first 3 pillars: institutions, human capital, and infrastructure, while Kenya scores higher in remaining two input pillars, and both output pillars.

Another pair of countries plotted close to each other and belonging to different clusters is Dominican Republic (87th overall rank) & Morocco (76th). Both of the countries are characterized by higher innovation outputs than inputs, with Morocco being on 65th position in innovation efficiency and Dominican Republic on the 71st. It is also interesting to note that Morocco is a representant of lower-middle income group, while Dominican Republic belongs to upper-middle. Morocco scores higher in 3 innovation inputs sub-pillars: institutions, human capital, and infrastructure, as well as, knowledge & technology outputs. Dominican Republic performs better in market and business sophistication, and is one rank higher in creative outputs.

It is worth coming back to the bases of hierarchical clustering and remember that the algorithm connects each observation to the one, or to a pair, of observations that are the most similar based on the set similarity measure, and connectivity method. The dendrogram, shown as figure 23 in the Appendix, will be useful in this case. It can be seen that members of the red cluster were connected: Morocco to Uruguay, and Jordan to Tunisia. While the members of green cluster were connected: Dominican Republic to Indonesia, and Kenya to a pair consisting of Armenia and Tanzania. It could be beneficial to look at characteristics of those countries, however some patterns can already be seen from comparison of the two pairs mentioned above. Both members of the red cluster (Morocco & Jordan) scored higher in the first 3 innovation inputs sub-pillars, which also had more spread values based on the first exploration of mean scores probably differentiating the observations more, while members of green cluster had better performance in both market and business sophistication. As there has been only one rank difference in both cases between creative outputs, it cannot be analyzed how does it affect the formation clusters.

6.5 Characteristics of top 3 economies

Lastly, as the main purpose of the analysis is discovering what are the characteristics of the most innovative economies, the top 3 countries from the overall ranking will be analyzed. They all have been on the first 3 ranks for the past 2 years. Figure 22 in the Appendix presents the top 10 most innovative economies in the ranking and their changes in position within the last 5 years. Quite interesting to note that Switzerland has been ranked first the whole time, while other economies change their innovative performance.

Sweden is ranked 3rd in the overall ranking, as well as, in both innovation inputs and outputs, giving the country 10th place in efficiency ratio. It is ranked as 9th in institutions pillar, with the main strength in the indicator measuring rule of law, leading overall ranking, and the main weakness in cost of redundancy, on the 55th place. It scored 11th place in education sub-pillar, with 4th rank for expenditure on education as a percent of GDP (7.7%), and pupil/teacher ratio (52nd place) as the main weakness. Sweden has 29th place in tertiary education sub-pillar (the lowest from all sub-pillars). 62.3% of population is enrolled to tertiary education, placing Sweden on 38th position, with 26% of graduates in science and engineering (25th place). Despite low scores in tertiary education, Sweden is ranked 2nd in sub-pillar related to knowledge workers, led by 3rd place for 70.3% of firms offering formal training, and 5th place for knowledge-intensive employment.

The overall business sophistication pillar places the economy as 5th, with 4th rank for innovation linkages sub-pillar, mostly caused by 4th rank joint ventures/strategic alliances, but also 10th place for University/industry research collaboration. Sweden has 6th rank for research and development sub-pillar, with 3rd place for researchers hired full-time (per million citizens), and 4th place for gross expenditure on R&D as a percentage of GDP. Looking at the scores related to researchers and knowledge workers, it can be seen that low values of indicators related to tertiary education, do not affect the employment of qualified workers, or it could be an indication that if someone pursues a career as researcher, it would be quite easy to find employment. It also does not correspond to so-called European paradox, as the innovation inputs on science are actually worse than outputs in this case. Additionally, Sweden also scores high in other innovation outputs, not directly related to science. Its 4th place in knowledge & technology outputs is mostly driven by 1st rank for the number of international patent applications, and for intellectual property receipts. Sweden is also 6th for creative outputs with 3rd ranks for both online creativity and ICTs & organizational model creation.

Another strength of Sweden is 3rd place in infrastructure pillar, mainly driven by 6th rank for ICT use, 3rd position for logistics performance, 5th for environmental performance, and 7th for electricity output. This highlights emphasis of the country on sustainability. The main weaknesses of Sweden, considering all indicators is 104th place for FDI net inflows and 70th for the ease of getting credit. With 55th rank for R&D financed by abroad, it appears that domestic firms have more important role in Sweden than foreign investors. Additionally, Sweden ranks 53rd for high-tech imports but makes up for this score with exports on the 22nd, and high % medium-high-tech manufactures on 13th, suggesting that the domestic market is quite well-developed and Sweden can produce technologies for its own use.

Netherlands is placed second in the overall ranking, as well as, the innovation outputs. It is situated on the 4th place in efficiency ratio and 9th within innovation inputs. Netherlands is also the most interesting case in the top 3 because it has been on the 9th rank in 2016, changing to 3rd in 2017, and recently taking the 2nd place. I have attempted to analyze what was the cause of these changes but there is a limitation caused by the fact that GII framework has been constantly changed and improved, and there has been a significant revision in 2017. Therefore, from the whole framework, only 98 variables were corresponding and did not include any missing values for the last 3 years.

	Netherlands18	Netherlands17	Netherlands16
Global_Innovation_Index	2	3	9
Efficiency_Ratio	4	4	20
Input_Subindex	9	9	12
Output_Subindex	2	2	9

It can be seen that the improvements over time for Netherlands were both in innovation inputs and outputs, and the ranks have not changed since the last year. The shift from 9th to 3rd place could also be caused by changes in GII framework but in the following section, I will try to look into main changes within indicators that have data for all 3 years.

	Netherlands18	Netherlands17	Netherlands16
Foreign direct investment, net outflows	1	1	118
Knowledge diffusion	2	2	114
Gross capital formation	79	87	99
Graduates in science and engineering	82	88	88
Ease of getting credit	88	72	69

As the next step, indicators were sorted in descending order by the values for 2016, as it is known that this is the year with the lowest rank for Netherlands. It becomes clear from the table that there is something wrong with FDI net outflows value, which also affects the score for the whole knowledge diffusion sub-pillar, because it would be almost impossible to change country’s position from the end of the ranking into the lead within just one year. Indeed, looking at the reports from 2016 and 2017, it is explained that Netherlands lost its competitive position as a result of large fluctuations in selected data. FDI net outflows have been identified as highly volatile in previous GII editions, partly driving the fall in the ranking for Netherlands. Also, Netherlands lacked data for some new variables: IP receipts and ICT services exports. Thus, the shift in rankings over the last 3 years is mainly caused by issues with collecting data and measuring innovation properly.

The last country to analyze is Switzerland - the leader of Global Innovation Index ranking since 2011. Switzerland is also placed first in both innovation outputs and efficiency ratio, second in innovation inputs. It is also a leader in both innovation outputs pillars.

The economy scored 11th place in institutions pillar, with the 2nd place in political environment, 5th in regulatory environment, and 44th for business environment. The main weaknesses indicated within this pillar is 32nd rank for the cost of redundancy, 42nd for the ease of resolving insolvency, and 59th position for the ease of starting a business. Switzerland also has 61st rank for ease of getting credit, and 92nd in ease of protecting minority investors. Thus, it can be concluded that Switzerland does not have a friendly environment for entrepreneurs, as they are required to go through a series of procedures, obtain additional licenses or permits to start a business, and then there is a risk of losing capital in case of insolvency.

The country is ranked on second place in the R&D sub-pillar, indicating that it is a favorable setting for innovation. The country is placed on the 5th rank for human capital and research pillar, with 32nd place for education. The low rank can be explained by 50th place for expenditure on education as a percent of GDP (5.1%), and school life expectancy of 16.2 years, placing Switzerland on 29th position. In addition, only 57.9% of population has tertiary enrolment (43rd overall rank), from which 24.4% of graduates have degrees in science & engineering (32nd place).

Switzerland is ranked on the 8 place for infrastructure pillar, with the main weakness lying in 30th rank for ICTs. The main reason for that is 64th place for government's online services, and 70th for e-participation, which is used to measure the provision of information by governments to citizens (e-information sharing), interaction with stakeholders (e-consultation), and engagement in decision-making processes. This indicator shows insights into how different countries use online tools to promote interaction between citizens, as well as, citizens and government.

Switzerland ranks the highest (4th) in business sophistication from all input pillars. Driven mostly by first place for University/industry research collaboration. The country also scores high for all indicators related to patents, leading in patent families filed in at least two offices, international patent applications, and IP receipts. Together with 2nd place for scientific & technical articles, high & medium-high-tech manufactures, 3rd place for computer software spending, and 5th for ICT services imports, it can be seen that Switzerland is focused on technologies and firms operating there protect their knowledge base. High values for expenditure on software and import of ICT services can be caused by the country's efforts to make up for what Switzerland lacks in infrastructure.

However, the leadership undertaken by Switzerland for the last 5 years has been criticized based on how much the ranking is a result of innovativeness, and how much it is based on international businesses being set up with headquarters in Switzerland for tax purposes (Skillicorn, 2016). As it is not the only critique of the innovation measurement and indicators included in the dataset, the following section will include discussion, also based on issues encountered while working with this particular dataset.

7. Discussion

To construct the ranking, the analysts rely on a series of numerical datasets that are currently available to compare countries. Direct measures, which quantify innovation outputs are still extremely limited and there are no official statistics on the amount of innovative activity measured by the number of new products, processes, or other innovations for any given country. Thus, it needs to be decided which available indicators represent the "innovativeness". There has been a critique of putting too much emphasis on macroeconomic and socio-political data, and too little on achieving innovation. There is a number of other indices that are focused on general economic health and strength. The authors justify it that underlying climate and infrastructure are vital for innovation but they have admitted that some measures are not specifically related to a country's ability to innovate. The report states that measuring innovation remains difficult, hence a great emphasis is placed on measuring the climate and infrastructure for innovation, and on assessing related outcomes. It is also said that measuring innovation linkages adequately remains challenging, if not impossible, based on the existing innovation metrics (Skillicorn, 2016)(Global Innovation Index, 2018). This issue is related to difficulty of capturing open innovation. Regions that promote knowledge diffusion, might even be ranked lower because good measures for it do not exist.

Most measures struggle to appropriately capture innovation outputs of a wider spectrum of innovation actors, such as public entities or services. For instance, innovation surveys that succeed in measuring innovation activities, fail to provide a reliable cross-country innovation output performance, and often are not applicable to developing countries where innovation is informal. In many cases, innovation is actually driven by challenges existing in a local market which need a novel solution, for example in African countries with neither strong financial institutions nor widespread and affordable ICT access (Skillicorn, 2016). Therefore, other forms of innovation (such as frugal in this case) should also be incorporated in assessing the innovativeness of countries. Besides, the outputs are mostly measured in a way the more the better, which do not distinguish innovations from imitations, and emphasized quantity over quality.

Cultural outputs are criticized for only remotely relating to innovation. In the case of printing and publishing, it actually considers an industry, which is in decline because of innovation in digital (Skillicorn, 2016). For instance, in case of Singapore, the country has different focus, shifted more towards high-tech, therefore it scores low in all the indicators related to culture and media, which is one of the reasons the country is considered an underperformer. So, it does not necessarily mean that the publishing industry is not developed but the indicator does not include digital media, which could be dominant for some developed economies.

There is also an indicator related to consumers and advertisers spending data for media & entertainment that shifts for example Japan's performance in cultural outputs. However, this indicator is not even related to innovation but also might not reflect the national features, as some consumers may choose foreign music, films, games etc.

Some indicators are based on qualitative data collected through answers to questionnaires on a firm level, which is related to a potential risk of different understanding of innovation between firms. The interpretation of results depends on how innovation was defined in the survey beforehand, and whether or not the firms understand what is meant by innovative activities. It is said that even if the questions in survey are intended to be the same, differences in language, culture, or context can lead relatively similar firms to give diverse answers (Fagerberg, 2013). It is a serious threat, as it has even led to the change of definition of innovation in Oslo Manual. Therefore, the outcomes of questions related to assessing the degree to which ICTs enable new business or organizational models can be biased based diverse interpretations or subjectivity in assessment.

Another issue of international comparison of innovativeness, based on various indicators is related to the definition of innovation that the analysis is based upon. For example, in the newest version of Oslo Manual (2018), the innovation is defined based on only process and product innovations, which comes back to the definition included in the manual from 1997. Oslo Manual is an internationally recognized guide for collecting and using data on innovation, therefore each time the definition of what is considered as innovation is updated, as a consequence researchers responsible for constructing various frameworks for innovation measurement also need to implement some updates. Firstly, it was argued that activities of firms, which are related to innovation are more than product and process innovations, and therefore, the definition should be broader (Bloch, 2007). However, when the theory met practice, it turned out that managers found it difficult to differentiate between organizational and process innovations, which caused disorientation and complications in making comparisons between various analyzed units (Oslo Manual, 2018). The reason why product and process innovation gain the most attention and focus could be that they are useful for the analysis of some issues. However, it is said that other kinds of innovations should not be ignored, as many innovations that helped the United States to outperform other capitalist economies, during the first half of the 20th century, were based on the organizational innovations (Fagerberg, 2013).

With the changes in definitions, there needs to be an update of indicators and different kind of data needs to be collected, which causes further problems. As it could be seen on an example of Netherlands, even the state for the last 3 years cannot be compared, because the framework is constantly updated, which causes both many cases of missing data, and difficulties in analysis over time, as some indicators are replaced by new measures.

In the post-war period, the strong tendency towards internationalization of trade, capital, and production has developed. Discussions have been raised that nations as economic subjects are no longer in use. But the process of internationalization has not erased the individual national patterns of specialization in production or international trade. There can be several reasons why interactions between users and producers which belong to the same national system might work more efficiently. One part of the explanation is short geographical distance, together with sharing common language and cultural proximity. It is interesting to note, that for example Scandinavian firms tend to regard all the Nordic countries as their home market, as they have a lot in common in terms of culture and social organization (Lundvall, 2016). Since differences in technological capabilities between various regions are very broad, one can doubt the usefulness of comparing countries with completely diverse characteristics. Comparisons become more significant if they are carried out between more similar national systems of innovation, for example between Nordic countries (Archibugi et al., 2009). The idea behind using unsupervised learning techniques on this dataset is also based on similarity measures. It is easier to find what are the common features of economies grouped together, than to find any patterns between observations, which have completely diverse features.

Innovation can be defined as a process, including generation of new ideas, design, development, industrialization, commercialization, transmission, and diffusion. On the contrary to what is assumed in the linear model of innovation, these above-mentioned phases are not sequential but interrelated. Expertise and competences within different phases are not equally distributed within the economic space. Therefore, some countries show a greater capability to generate new ideas, for example because there are many reliable public research centers, good universities or efficient industrial labs. Other countries are more capable of exploiting them commercially, while some are more inclined to absorb knowledge acquired externally and diffuse it internally, although it is said that most diffusion processes result in the generation of incremental, adaptive innovations (Archibugi et al., 2009).

While researchers conduct analysis on a country level a key issue with methodological assumptions is that countries are made of different areas and regions, which are far from being homogeneous. Therefore, using a single indicator to capture the overall technological capabilities of such differentiated entities consists of several simplifications. It is common within macroeconomic analysis to encounter this type of simplifications, the gross domestic product is used on a daily basis, even if its real economic meaning is often being questioned because the indicator aggregates very heterogeneous phenomena in a single figure. The other example could be the unemployment rate. While considering the aggregate rate of unemployment, the fact that some regions can be characterized by full employment, while in others unemployment rate can be far higher than the national average, is disregarded, thus analysts lose some proportion of valuable information assuming the homogeneity on a country level. There are also significant differences across regions, industries, and companies within the same country regarding the measurement of technological capabilities. As a result, the inter-country comparisons are possible based on the implicit assumption that a national system of innovation is capable, in some way, to distribute knowledge across the whole country (Archibugi et al., 2009).

Let’s take into consideration the United States - it leads the ranking for state of cluster development, scores for top 3 universities, and biggest R&D spending by the top 3 companies. However, this state does not represent the whole country. It can be seen that there are some specific companies or cities that shift these indicators, and the data would not be homogenous if the same measurements were compared between regions. The big companies, located in high-tech clusters are the main drivers in protecting their IP rights, R&D investments, and University/industry research collaboration. But many other countries are divided into different regions that have separate legislation and even big cities have mayors that make decisions on a local level. Therefore, there might be a big variation in the focus of each region, as some might rely on specific industries, sectors, or natural resources. Moreover, tertiary education institutions, as well as, big firms tend to be located in cities, so if the state for cities and countryside are compared, there would be significant differences.

Innovation processes differ significantly between sectors. Aspects such as organizational structures and institutional factors, availability of skills, linkages and access to knowledge, rate of technological change, or the presence of high-technology sectors have an important role in the capacity of a country to transform R&D into innovations. The economic structure of a country plays an important role in the processes of generation and assimilation of innovation. Some sectors are characterized by rapid change and radical innovations, while others involve smaller, incremental changes. For example, in high-technology sectors R&D plays a central role in innovation activities, on the contrary to other sectors, which rely to a greater extent on the adoption of technology and knowledge. Economies, which rely on technologically advanced sub-sectors have greater tendency to note higher rates of innovation. Those differences between innovation activities also place diverse demands on the organizational structure of firms within a sector or institutional factors, such as regulations or intellectual property rights, which affect designing policies on regional or country levels (Oslo Manual, 2005)(Bilbao-Osorio & Rodríguez-Pose, 2004). Those countries or sectors, characterized by bigger protectionism, will score higher in innovation outputs. R&D and patent-based measures also characterize closed innovation. Even though protecting IPs does not foster the diffusion of knowledge, economies that apply for a lot of patents, trademarks, or industrial designs will rank higher in innovation. Those indicators are also easy to measure, so it is not likely that they will be replaced, and therefore countries that promote knowledge exchange and do not popularize protecting technologies will remain less competitive in terms of knowledge & technology outputs.

However, those measures have been criticized as many innovations are not patented, and patents are awarded not for innovations but inventions, so using a patent-based measure of innovation may lead to a biased picture (Fagerberg, 2013), which is also not in line with a defining characteristic of innovation that the new idea, method, model or prototype needs to be implemented. This is another issue with measuring innovation - the indicators that are available and quite easy to collect, might not reflect an actual state within the country.

When the composite indicators are obtained as the arithmetic mean of single statistics, one unit can be substituted by a unit of another indicator and vice versa. Considering the fundamental differences between the aspects measured by different indicators, this assumption has been criticized that the various components need to be weighted. However assigning weight to each indicator, provides a subjective value judgment (Archibugi et al., 2009). The first issue that arises from practical analysis of the dataset is related to what is stored in different variables and the fact that if all values can be substitutes. Some countries will be ranked high not for being actually innovative but for having good infrastructure or institutional set-up, as from the analysis of clusters it turns out that the first 3 pillars were the biggest determinants of differences in innovation inputs. If a country is lacking in one area, it is possible to improve its competitive performance in the ranking by compensating it with increase in another one, which may not improve the situation in reality, just affecting higher rank in the comparison of economies. It is also easy to affect the scores by specializing in one area, for instance by improving all the scores related to knowledge creation or filling many patent applications. However, this does not reflect the quality of innovation.

But various elements of aggregated indicators are only treated as substitutes in terms of statistical importance and the assumption does not have anything to do with reality. If a country is lacking in one area of development, it is possible to improve this economy’s performance in the ranking by compensating it with increase in other area, but it might not affect the country in a positive way. Thus, the aim of decision-makers should not be to increase the value of indicators but to improve the economic and social conditions these indicators are expected to capture (Archibugi et al., 2009).

8. Conclusion

From the way aggregated indicators are constructed, there is not a single solution on how to be more innovative. Even comparison of the top 3 countries in a ranking mostly shows drawbacks of this kind of data, where each pillar and variable is weighted the same way and the ranking is constructed simply on the average between innovation inputs and outputs. Each country in the top 3 has different strengths and weaknesses, for instance Switzerland leads overall ranking in both output pillars and innovation efficiency ratio. Actually, all 3 countries have the same competitive positions in innovation outputs, as in the overall ranking. The main difference lays within input pillars or rather in the climate for innovation. Netherlands leads in business sophistication and Sweden is 3rd in infrastructure but except for that, the 3 economies do not score the highest in other input pillars. In fact, Singapore leads the ranking for this sub-index but because of being an underachiever, does not belong to the top 10 in the overall ranking.

Some issues with the data can be seen from a simple correlation matrix between the main indices. The strongest correlation coefficient is between the main ranking and innovation outputs, followed by innovation inputs. The correlation coefficient between positions in the Global Innovation Index and Innovation Efficiency Ratio is much lower, so it can be concluded that the performance of countries is determined more by how much inputs or outputs the economies create, rather than measuring how efficiently the inputs are being used. Thus, it is possible that some nations will be highly ranked, based mostly on good environment for innovation that they have or high investments in innovation, not necessarily using them efficiently. The efficiency ratio itself has stronger positive correlation with innovation outputs than inputs, meaning that countries should focus on producing innovations rather than just investing in them to be ranked higher in this measure.

The main users of aggregated indicators are mass media, economists, politicians, and managers. For mass media it is just a form of shaping public opinion by the direct ranking of countries. Economists use the innovation indicators to analyze relationships between innovation and other economic phenomena, such as competitiveness, trade, growth or productivity. Policymakers and managers are said to be keen on reading and commenting on this data, but being less eager to guide their actions on the ground of those indicators. One of the reasons for that might be the realization that those rankings are far too aggregate to be connected to specific policies or strategies (Archibugi et al., 2009).

There are many definitions of national systems of innovation and different scholars distinguish diverse determinants of innovation. Indicators used in Global Innovation Index use the broad definition, including all the factors that influence innovation processes, because the exact determinants of innovation are not known systematically and in detail. Generally, most elements of the competitive environment that are described in the theoretical considerations of this project are included. However because such a broad perspective on national innovation systems has been taken, some measures show just the general state of economy. Thus, not only the main results of Global Innovation Index should be analyzed but also its specific components. First reason for that is discovering measures that are actually related to innovativeness. Secondly, low scores of some indicators do not necessarily mean disadvantages. For instance, if the cost of redundancy is high, it means that the employees are protected by labour laws or high results related to intellectual property may only reflect protectionist approach of a government, or characterize high density of high-tech companies that shift those indicators. There is an additional issue with the dataset, as there are only ranks stored for each observation, so in order to obtain more details and specific values, there is a need to look into reports.

Additionally, policy-makers or researchers should not only rely on nationwide data and consider differences between regions within the country. Especially if the ranks for cluster development are high, it means that there are regions in the country that can be much more innovative than others. There are also indicators that reflect average scores for the top 3 universities or firms, which should be developed taking into consideration more institutions in the future to reflect less biased picture of the whole country.

R&D in some sectors can also be linked to the country-specific institutional features, which form a system that either supports or prevents the accumulation and diffusion of knowledge between the scientific and industrial communities. Therefore, it becomes important for policy analysts to understand the nature of those country-specific factors that have influence on creating national technological advantage, including the competitive environment, the financial system, education, training, and basic research institutions (Carlsson, 2006). Thus, macroeconomic indicators related to more general climate for innovation may be useful in identifying obstacles to innovation and causal relationships. If for instance weak ICT infrastructure translates directly to lack of outputs related to technologies, or weaknesses in educational institutions result in low outputs related to publications or knowledge workers.

Global Innovation Index, disregarding all the limitations, can still be useful to benchmark different countries, take important decisions concerning international cooperation based on complementary competences, or to study convergence and contribution to common goals, for instance within the European Union. Additionally, it is a good source of identifying advantages or obstacles to innovation. Solely the analysis of a specific country profile, provides a good, however incomplete, overview on the economy. But if it is possible to obtain complementary data on a regional level, as well as, the main co-operators and competitors, it would give sufficient information on strengths and weaknesses of a country.

9. Reference List

Amoros, R. (2018). *How Much Countries Invest in Research & Development.*
[Online] Available at:
https://howmuch.net/articles/research-development-spending-by-country

Archibugi, D., Denni, M., & Filippetti, A. (2009). The technological capabilities of nations: The state of the art of synthetic indicators. *Technological Forecasting and Social Change*, *76*(7), 917-931.

Arundel, A., Lorenz, E., Lundvall, B. Å., & Valeyre, A. (2007). How Europe's economies learn: a comparison of work organization and innovation mode for the EU-15. *Industrial and corporate change*, *16*(6), 1175-1210.

Bilbao-Osorio, B., & Rodríguez-Pose, A. (2004). From R&D to innovation and economic growth in the EU. *Growth and Change*, *35*(4), 434-455.

Bloch, C. (2007). Assessing recent developments in innovation measurement: the third edition of the Oslo Manual. *Science and Public Policy*, *34*(1), 23-34.

Borrás, S. (2004). System of innovation theory and the European Union. *Science and Public Policy*, *31*(6), 425-433.

Cantwell, J. (2005). Innovation and competitiveness. *The Oxford handbook of innovation*, 543-567.

Carlsson, B. (2006). Internationalization of innovation systems: A survey of the literature. *Research policy*, *35*(1), 56-67.

Desjardins, J. (2018). *Innovators wanted: these countries spend the most on R&D.*, World Economic Forum.
[Online] Available at:
https://www.weforum.org/agenda/2018/12/how-much-countries-spend-on-r-d?fbclid=IwAR0Jr46Qzeva2T74uB9n8DSQggeqPcuJ-wS0R5uo6dOrY467YELoyBESuLA

Edquist, C. (2010). Systems of innovation perspectives and challenges. *African Journal of Science, Technology, Innovation and Development*, *2*(3), 14-45.

Fagerberg, J. (2013). Innovation - a New Guide. *Working Papers on Innovation Studies*, University of Oslo, Centre for Technology, Innovation and Culture

Fagerberg, J., & Godinho, M. M. (2004). Innovation and catching-up. Georgia Institute of Technology.

Fagerberg, J., Srholec, M., & Knell, M. (2007). The competitiveness of nations: Why some countries prosper while others fall behind. *World development*, *35*(10), 1595-1620.

Hayden, L., (2018). *Principal Component Analysis in R.*
[Online] Available at:
https://www.datacamp.com/community/tutorials/pca-analysis-r

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: a review. ACM computing surveys (CSUR), 31(3), 264-323.

Jaiswal, S. (2018). K-Means Clustering in R Tutorial
[Online] Available at:

https://www.datacamp.com/community/tutorials/k-means-clustering-r

Laufer, J. (2019). *Missing value visualization with tidyverse in R.*
[Online] Available at:
https://jenslaufer.com/data/analysis/visualize_missing_values_with_ggplot.html

Lundvall, B. Å. (2016). Innovation As An Interactive Process: From User–Producer Interaction To The National Systems Of Innovation. *The Learning Economy and the Economics of Hope*, 61.

Lundvall, B. Å. (2017). *The Learning Economy and the Economics of hope*. Anthem Press.

Organisation for Economic Co-operation and Development., & Statistical Office of the European Communities. (2005). *Oslo manual: Guidelines for collecting and interpreting innovation data.* Paris: Organisation for Economic Co-operation and Development.

OECD / Eurostat (2018), *Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation, 4th Edition*, The Measurement of Scientific, Technological and Innovation Activities, OECD Publishing, Paris/Eurostat, Luxembourg

Pathak, M. (2018). *Hierarchical Clustering in R*
[Online] Available at:
https://www.datacamp.com/community/tutorials/hierarchical-clustering-R

Patil, P. (2018). *What is Exploratory Data Analysis?*
[Online] Available at:
https://towardsdatascience.com/exploratory-data-analysis-8fc1cb20fd15

Rakas, M. & Hain, D. (2019). The state of innovation system research: What happens beneath the surface? *Research Policy*.
[Online] Available at:
https://doi.org/10.1016/j.respol.2019.04.011

Skillicorn, N. (2016). *Why I have a problem with the Global Innovation Index*.
[Online] Available at: https://www.ideatovalue.com/inno/nickskillicorn/2016/09/problem-global-innovation-index/

The Global Innovation Index (2017). *The Global Innovation Index.*
[Online] Available at: https://www.globalinnovationindex.org

The Global Innovation Index (2018). *The Global Innovation Index.*
[Online] Available at: https://www.globalinnovationindex.org

Wickham, H. (2014). Tidy data. *Journal of Statistical Software*, *59*(10), 1-23.

Yan, S., Xu, D., Zhang, B., Zhang, H. J., Yang, Q., & Lin, S., 2007. Graph embedding and extensions: A general framework for dimensionality reduction. *IEEE transactions on pattern analysis and machine intelligence*, 29(1), 40-51.

10. Appendix

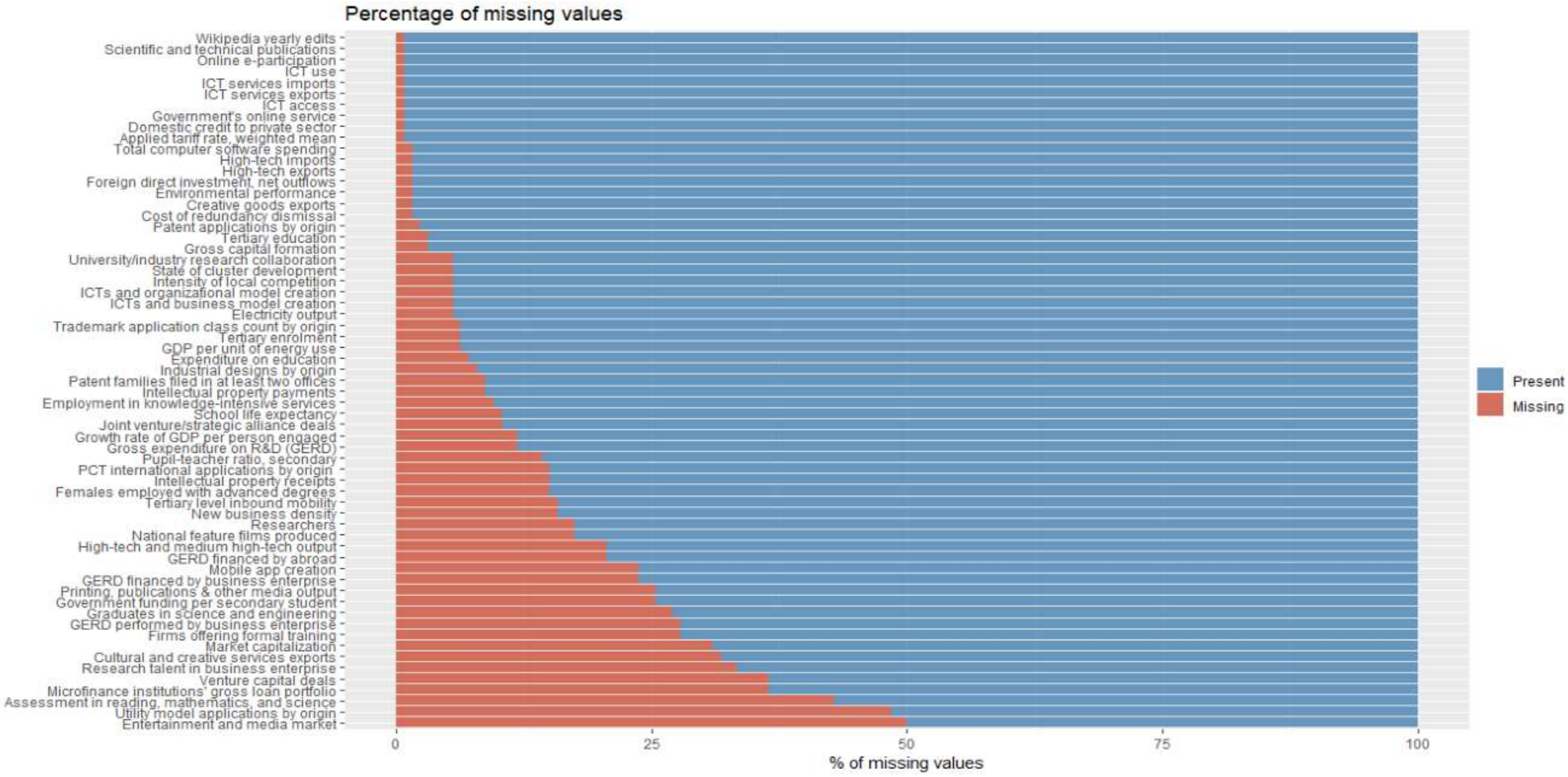

Figure 13: Percentage of missing values

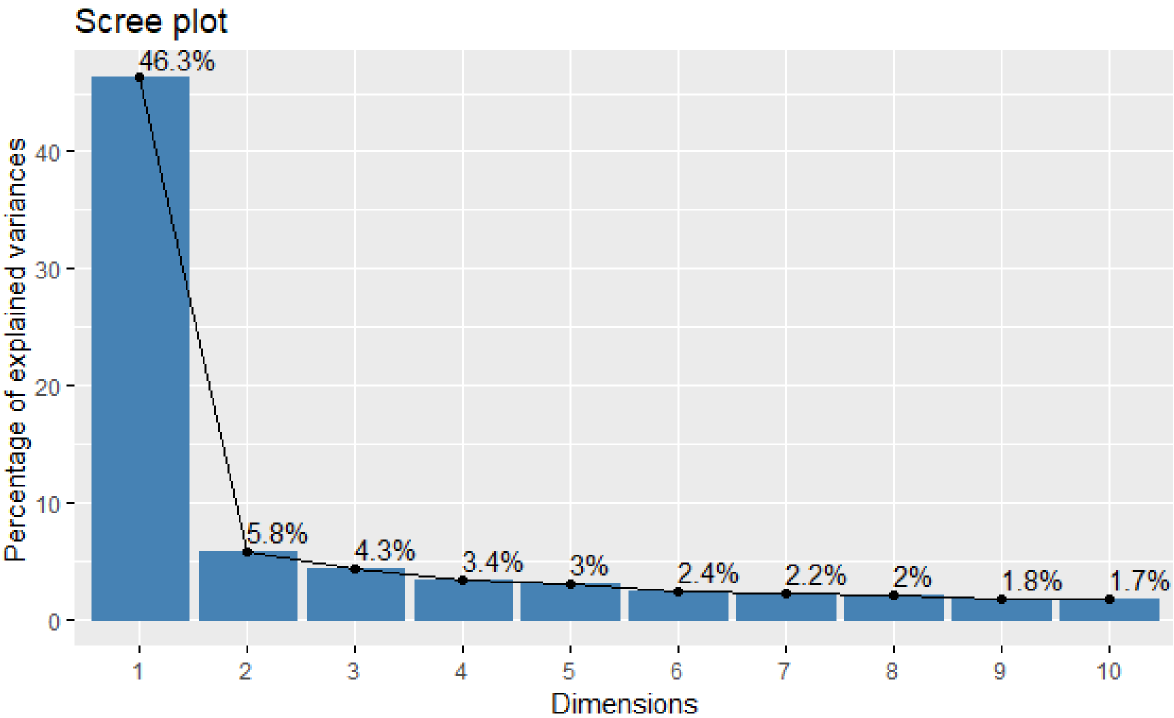

Figure 14: Scree plot for all indicators

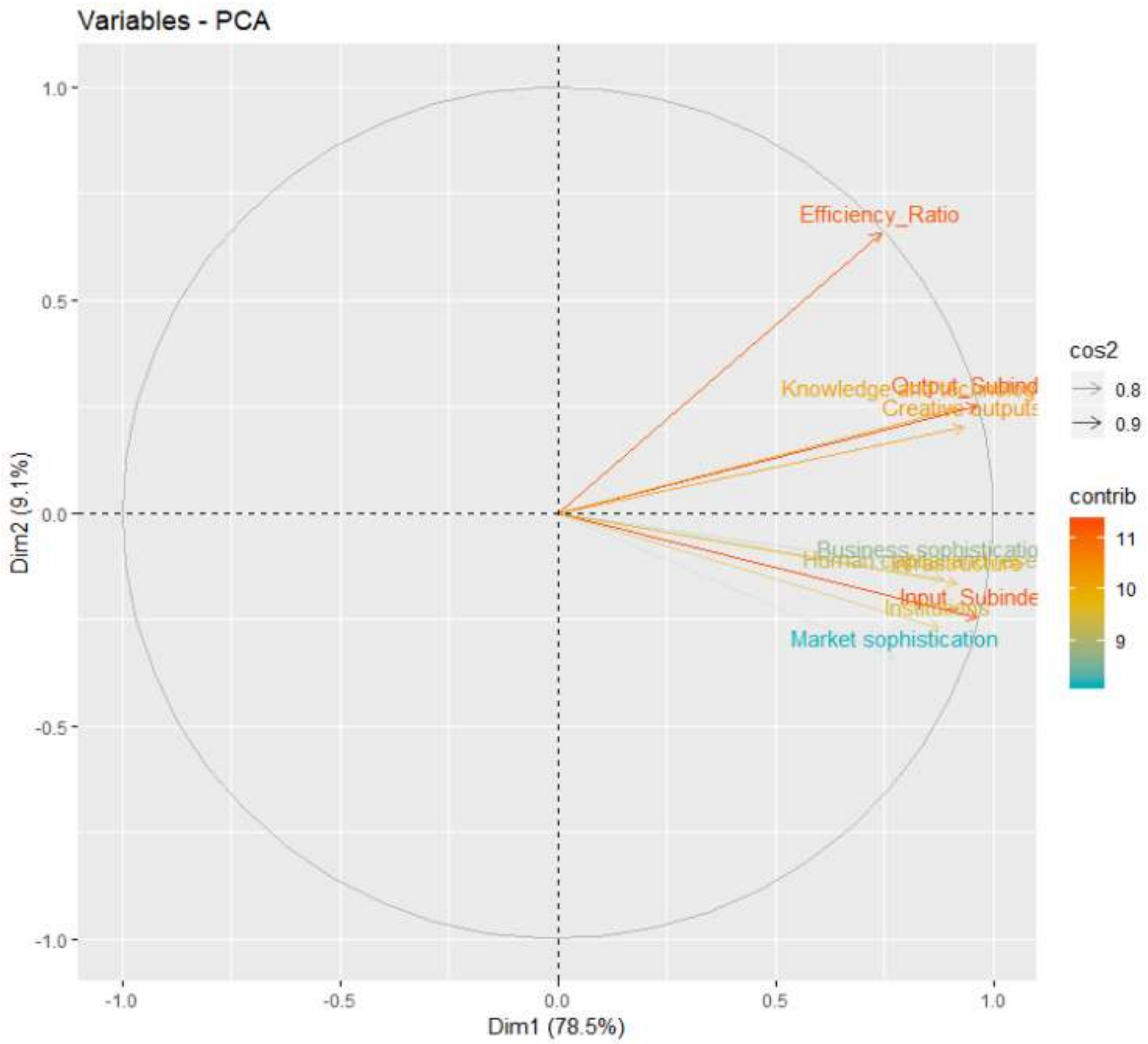

Figure 15: PCA biplot of chosen variables

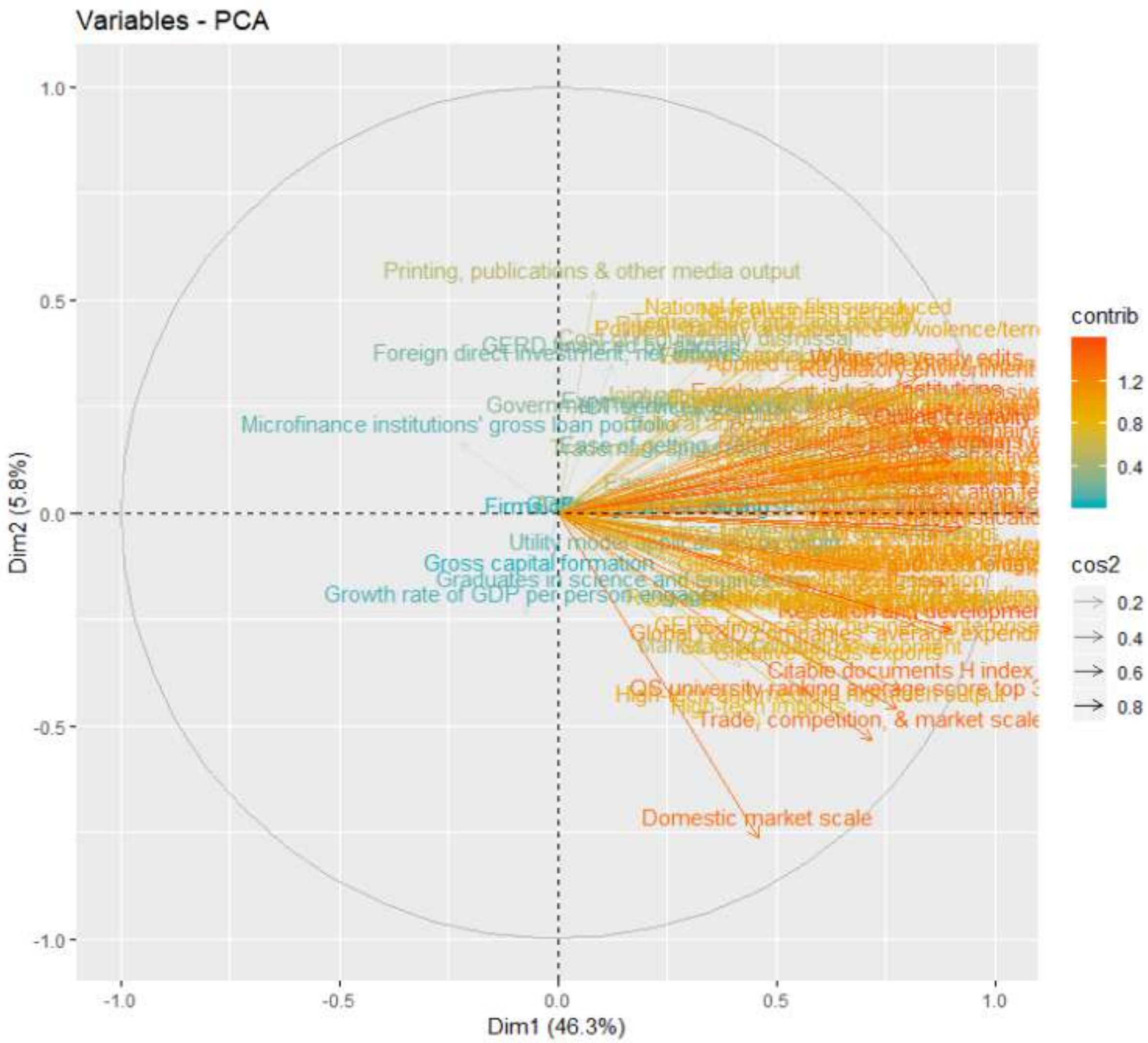

Figure 16: PCA for all indicators

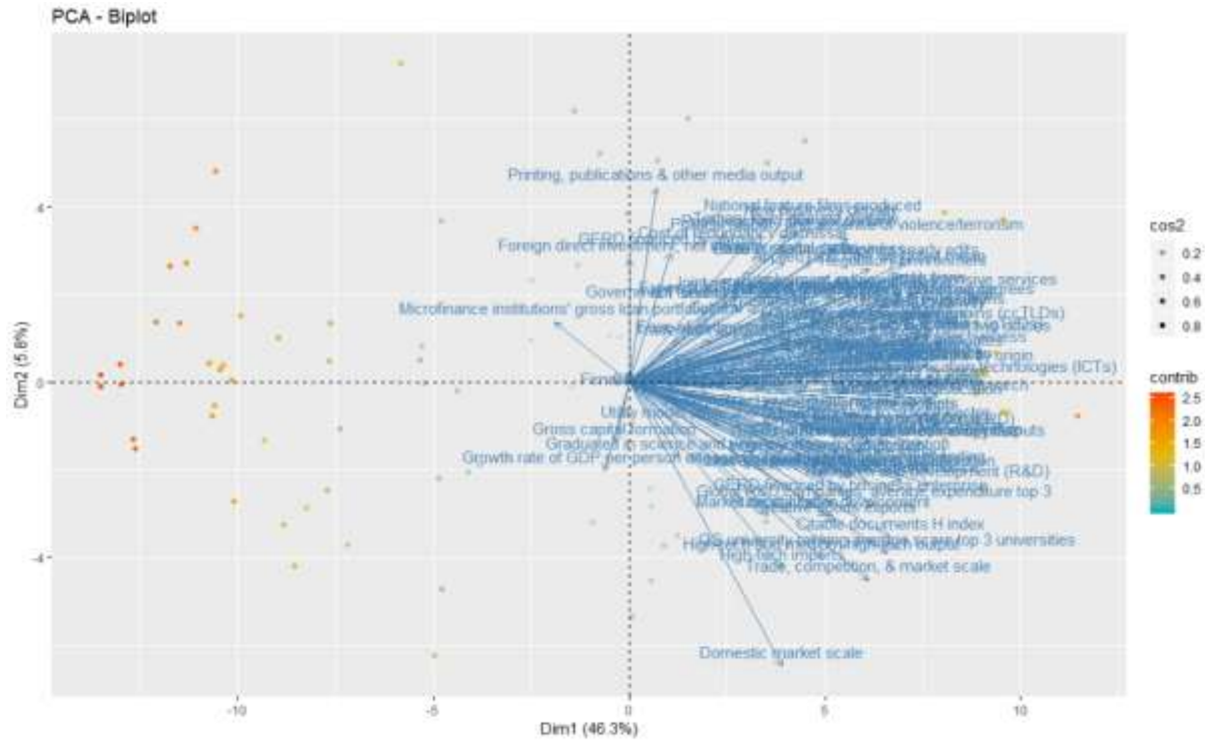

Figure 17: Biplot for all indicators

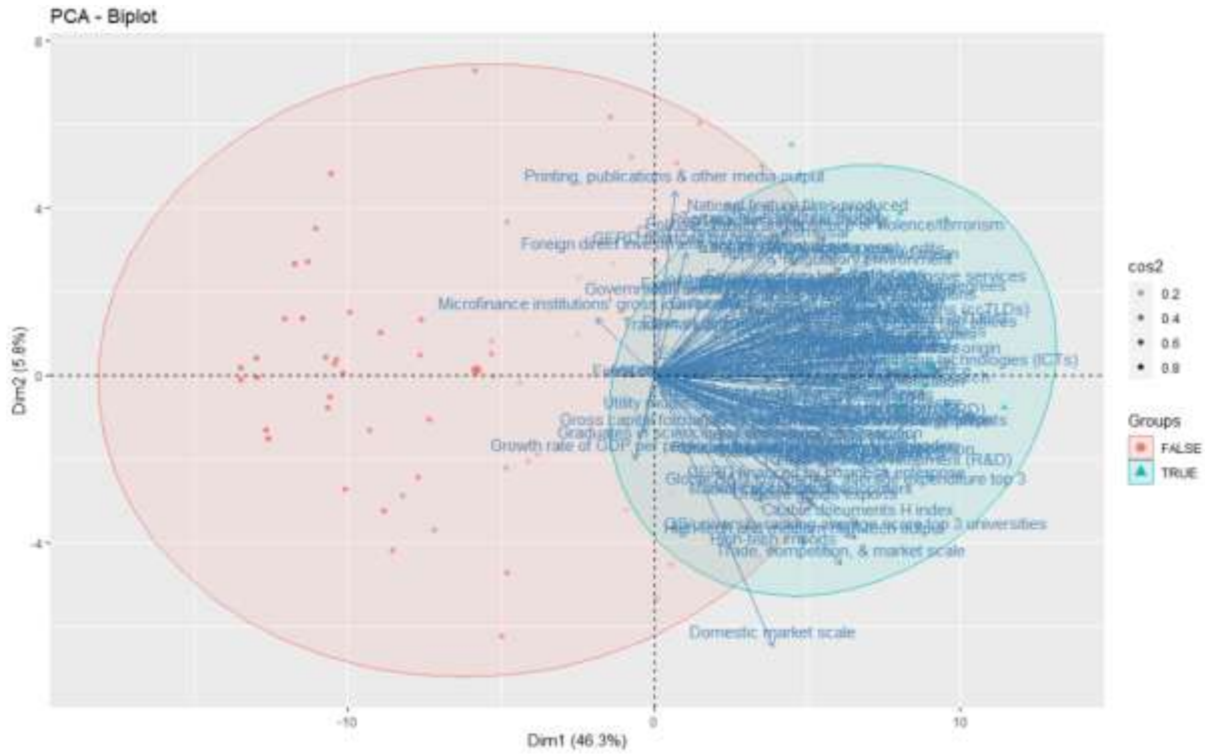

Figure 18: Biplot for all indicators grouped by ranks in GII

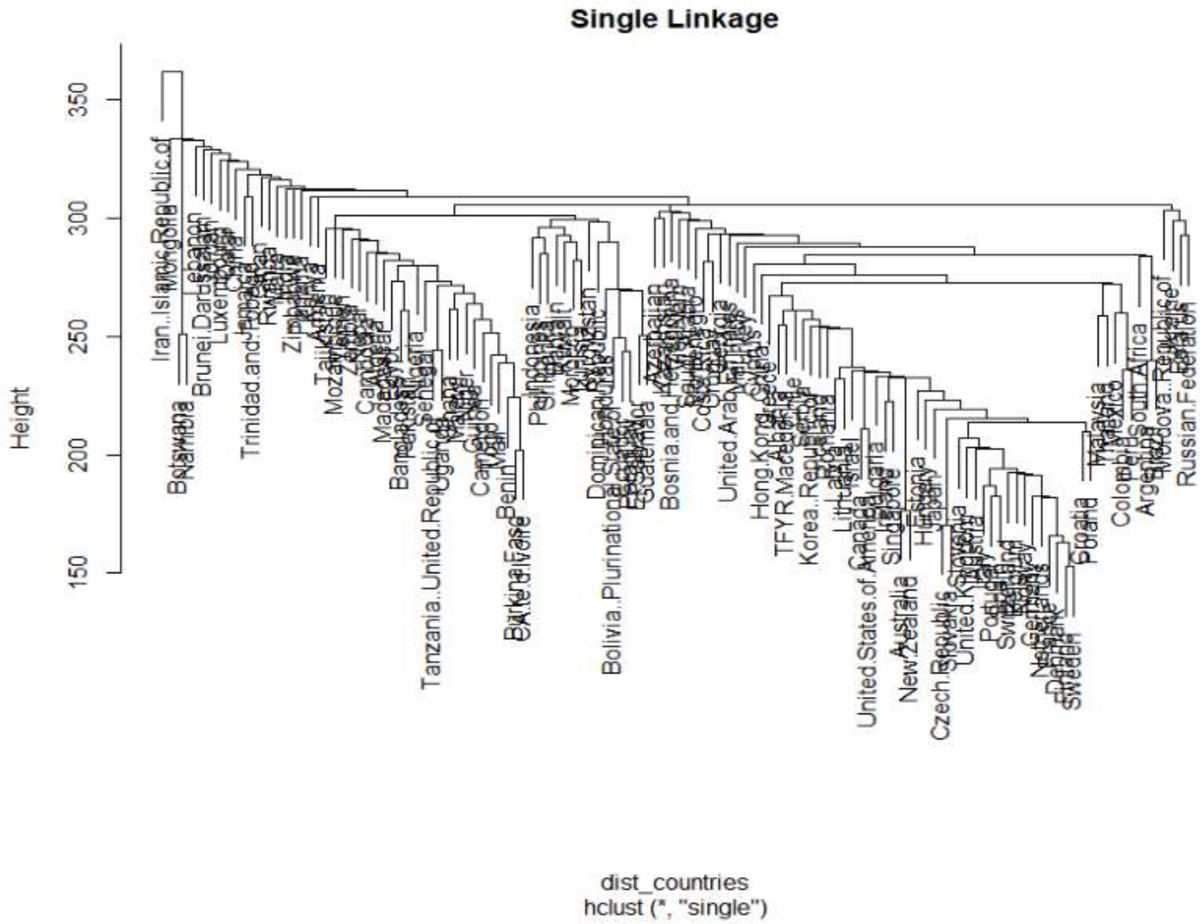

Figure 19: Clustering - single linkage

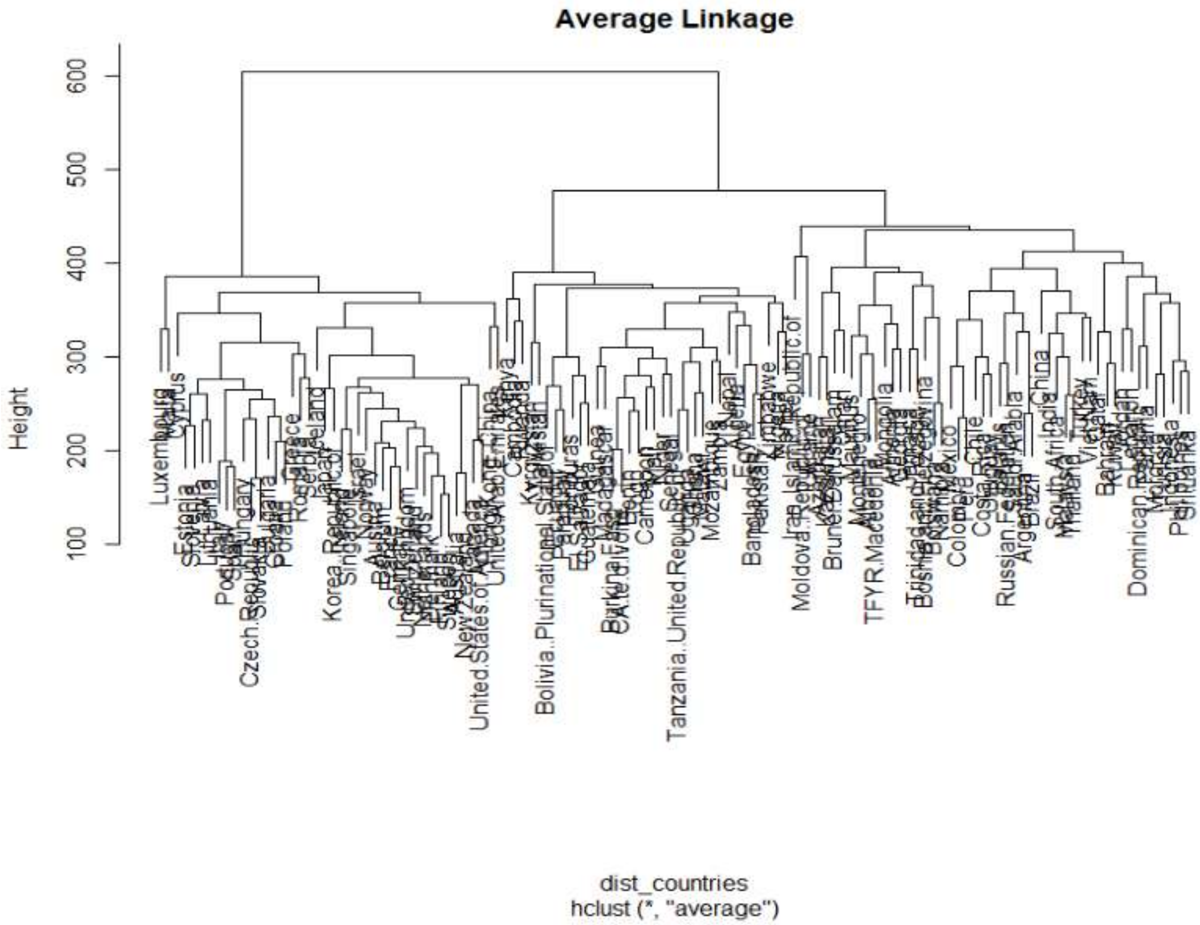

Figure 20: Clustering - average linkage

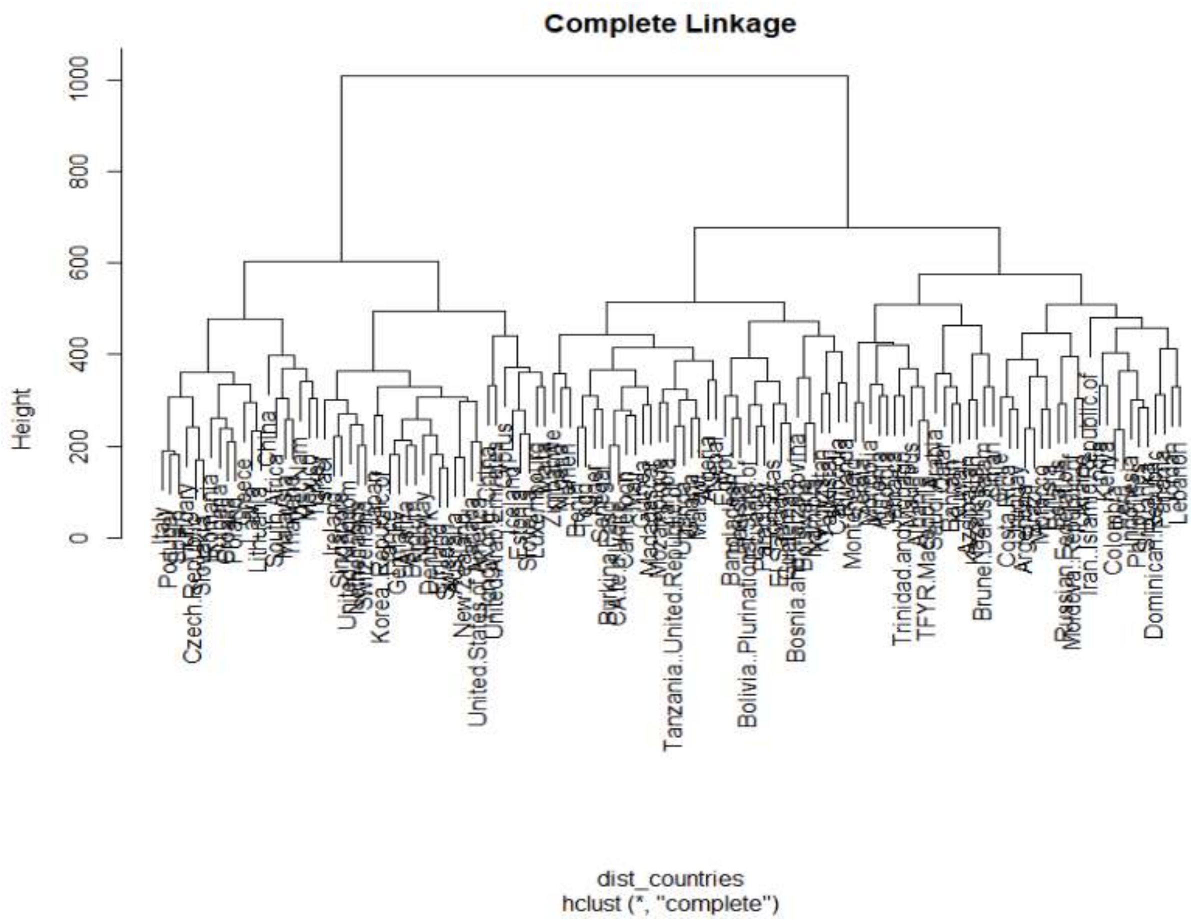

Figure 21: Clustering - complete linkage

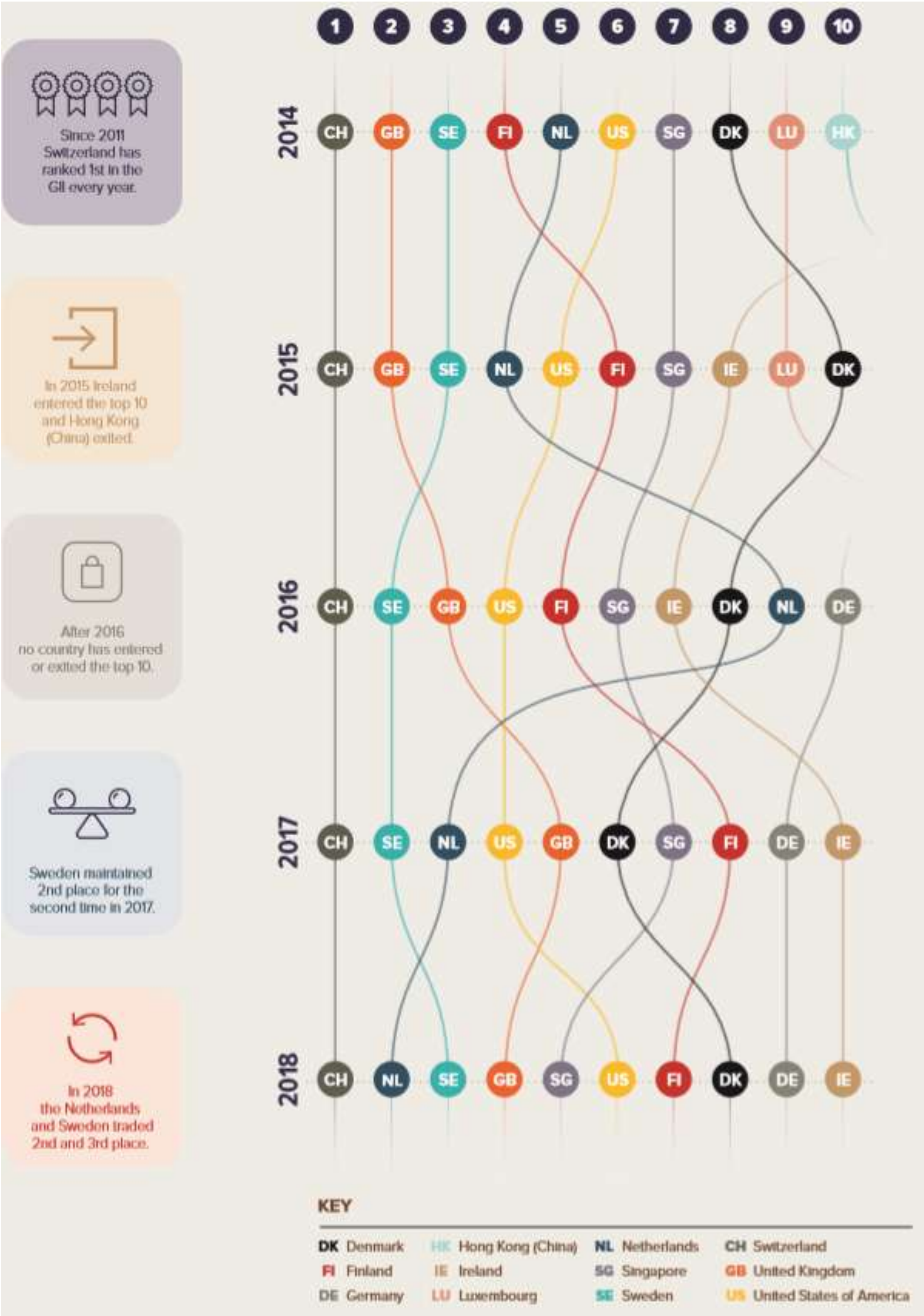

Figure 22: Changes in top 10 ranks (Global Innovation Index, 2018)

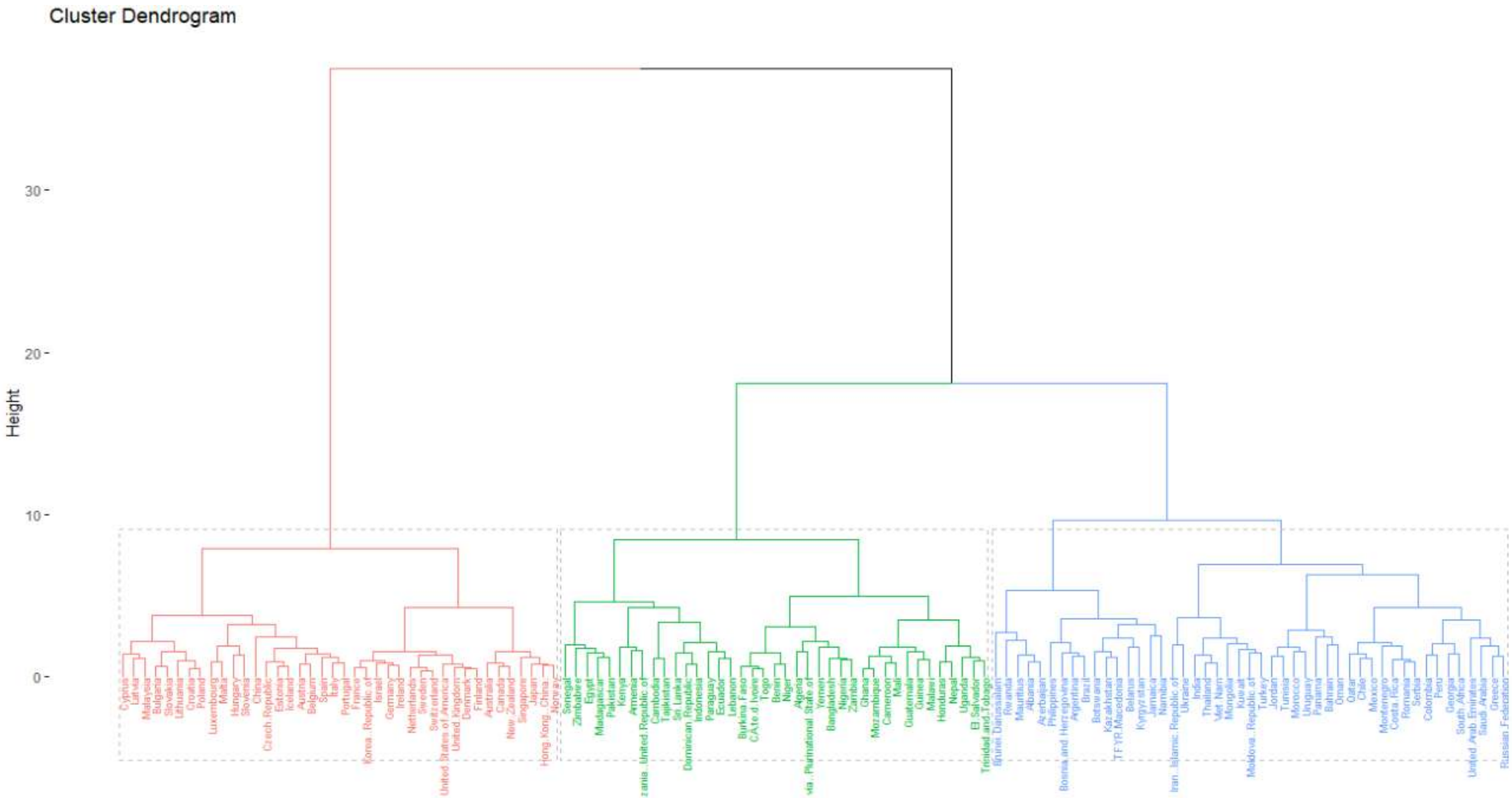

Figure 23: Dendrogram