Universidad de Deusto
University of Deusto

ON THE ROBUSTNESS OF COMPOSITE INDICATORS:

EVIDENCE FROM THE GLOBAL INNOVATION INDEX

Deusto Business School

Doctoral Dissertation

PhD programme in Business and Territorial Competitiveness, Innovation, and Sustainability

Doctoral candidate:

Khatab Ahmad Khalaf Alqararah
Deusto Business School
University of Deusto

Supervisors:

Dr. Jon Mikel Zabala-Iturriagagoitia
Deusto Business School
University of Deusto

Dr. Javier Barbero
Department of Economic Analysis
Autonomous University of Madrid

Donostia - San Sebastián (Spain), September 2023

ACKNOWLEDGMENTS

Pursuing a PhD degree was a lifelong ambition for me, and the journey to this remarkable milestone has been long and transformative. Along the winding path of my Ph.D. expedition, I have acquired the invaluable skills of scientific thinking, knowledge-seeking, and academic writing. Furthermore, receiving this esteemed degree from an institution as illustrious as Edisto Business School is a privilege beyond measure. I have invested numerous hours of hard work and faced many sleepless nights in pursuit of this goal. This achievement serves as a clear testament to the strength of my perseverance and determination.

At this juncture, I am deeply indebted to those whose support and guidance have propelled me forward. First and foremost, I extend my heartfelt gratitude to my supervisor, Dr. Jon Mikel Zabala-Iturriagagoitia. His continuous support and enthusiasm have been the guiding light throughout this academic odyssey. Dr. Zabala-Iturriagagoitia's positive outlook and understanding have been a constant source of motivation, for which I am truly grateful. I must also express my appreciation to Dr. Iñaki Peña Legazkue, the director of the CETIS doctoral program, whose support, firm guidance, and encouragement were instrumental in shaping the course of my research. To the members of the revision and defense committee, I extend my sincere thanks for their time and effort in evaluating my dissertation. Your insights and feedback have been invaluable in refining my work.

This arduous journey would not have been possible without the limitless love and support of dear individuals who stood by me. To my parents, Ahmad and Khawla, your continuous prayers and blessings have been my steadfast companions on this challenging journey. This dream would not have materialized without your love and support. To my beloved wife, Hanna, who believed in me from the very first step of this journey and has been a pillar of love and support throughout, I am forever grateful. To my precious daughter, Nura, your infectious smile has been a constant source of happiness and joy, providing me with the inspiration to persevere. Special thanks are also due to Dr. Ibrahim Alnafra for his invaluable assistance with Cluster Analysis and Mr. Amjad Kiwan for his dedicated support with Random Forests Analysis. Your expertise and responsiveness were instrumental in the success of my research.

In conclusion, this Ph.D. journey has been a transformative stage in my life, and it would not have been possible without the support and encouragement of these exceptional individuals and institutions. I am profoundly grateful for their contributions, and I look forward to continuing my academic and professional journey with the knowledge and experiences gained during this remarkable chapter.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PUBLICATIONS

Alqararah, K. (2023). Assessing the robustness of composite indicators: The case of Global Innovation Index. *Journal of Innovation and Entrepreneurship*, In press. https://doi.org/10.21203/rs.3.rs-1505335/v1

Alqararah, K., and Alnafrah, I. (under review). Innovation-driven Clustering for Better National Innovation Benchmarking. *Journal of Entrepreneurship and Public Policy*.

ABSTRACT

This dissertation introduces a methodology to assess the robustness of the Global Innovation Index (GII), by comparing the rankings provided in it with those achieved using alternative data-driven methodologies such as Data Envelopment Analysis (DEA) and Principal Component Analysis (PCA), using the data provided by the 2019 edition of the GII. With it, the dissertation aims to reduce the level of subjectivity in the construction of composite indicators (CIs) concerning weight generation and indicator aggregation. The dissertation relies on PCA as a weighting-aggregation scheme to reproduce the 21 sub-pillars of the GII before the application of the DEA Benefit-Of-the-Doubt approach to calculate the relative efficiency score for every country. By using the PCA-DEA model, a final ranking is produced for all countries. The paper uses Random Forests (RF) classification to examine the robustness of the new rank. The comparison between our rank and that of the GII suggests that the countries positioned at the top or the bottom of the GII rank are less sensitive toward the modification than those in the middle of the GII, the rank of which is not robust against the modification of the construction method. The PCA-DEA model introduced in this dissertation provides policymakers with an effective tool to monitor the performance of national innovation policies from the perspective of the relative performance of their respective countries.

Moreover, the dissertation employs a multi-dimensional innovation-driven clustering methodology to analyze the data provided by the 2019 edition of the GII. The K-means Cluster Analysis technique is applied to uncover and analyze distinct innovation patterns. As a result, it classifies 129 countries into four clusters: Specials, Advanced, Intermediates, and Primitives. Each cluster exhibits strengths and weaknesses in terms of innovation performance. Specials excel in the areas of institutions and knowledge commercialization, while the Advanced cluster demonstrates strengths in education and ICT-related services but shows weakness in patent commercialization. Intermediates show strengths in venture-capital and labor productivity but display weaknesses in R&D expenditure and higher education quality. Primitives exhibit strength in creative activities but suffer from weaknesses in digital skills, education, and training. Additionally, the study has identified 17 indicators that have negligible variance contributions across countries. The findings contribute to the field of innovation performance benchmarking by identifying appropriate benchmarking clusters and exploring learning opportunities and integration directions. To this end, it highlights the innovation structural differences among countries and provides tailored innovation policies.

Keywords: Composite Indicators. Data Envelopment Analysis. Innovation Performance. National Innovation System. Innovation Clustering. Innovation Benchmarking. Global innovation index.

Chapter 1
INTRODUCTION

Verily, the angels lower their wings for the seeker of knowledge- Prophet Muhammadﷺ .

INTRODUCTION

Innovation is an integral part of economic growth, and countries strive to improve their ability to innovate. It is a key factor for the success of any nation, leading to increased economic growth, job creation, and technological advances (Dziallas and Blind, 2019; Grupp, 2010; Yoon, 2017). At the country's level, innovation performance is largely dependent on a variety of factors such as government policies, access to education, infrastructure, and social determinants, among others. These factors serve as an explanation for the variation in a country's innovation performance (Zabala-Iturriagagoitia et al., 2007a; Dutta et al.,2019). Thus, a successful innovation performance measurement is necessary for a country to be competitive and prosperous in the global economy. However, innovation is characterized by its intricate and multifaceted nature, leading to diverse definitions within the literature. These definitions often vary depending on the specific socio-economic activities associated with different sectors, including business, public, and sustainable development sectors.

This common understanding is vital for introducing new indicators that effectively capture the interconnectedness and interactions among various actors involved in the innovation process. These indicators serve crucial purposes, such as evaluating and monitoring existing innovation policies, as well as informing the formulation of new policies (Gault, 2018). The latent nature of innovation has spurred the scientific community and stakeholders to devote considerable efforts toward constructing Composite indices (CIs) that seek to measure innovation at the national level. Notable examples of such indices include the European Innovation Scoreboard (EIS), specifically designed for European Union (EU) member states, and the Global Innovation Index (GII), which serves as an international benchmark. These composite indices are designed to capture and quantify the multifaceted dimensions of innovation, providing valuable insights into a country's overall innovation performance.

1.1. RESEARCH RATIONAL

1.1.1. Composite Indicators for Innovativeness Measurement

The need for innovation and knowledge-driven growth is no longer exclusively for developed countries, as developing countries are also acting to design policies aiming to enhance their innovation performance to achieve better economic and environmental growth (Broughel and Thierer, 2019; Nikolaidis et al., 2013). For a long time, continuous efforts have been deployed to construct CIs such as GII and EIS as a tool for measuring national innovation systems' performance (Bandura, 2011; Dutta et al., 2022; Edquist et al., 2018; Corrente et al., 2021; Alnafrah, 2021). Innovation, according to the Oslo Manual (OECD/Eurostat, 2018, p. 20) is defined as 'new or improved product or business process (or combination thereof) that differs significantly from the firm's previous products or business processes and that has been introduced on the market or brought into use by the firm'. In this regard, national innovation systems represent the combination and harmony of several facilitators and parties such as institutional structures, infrastructure, or any supporting activities and policies that orchestrate creating and facilitating an appropriate environment to foster innovation (Lundvall, 2007).

CIs represent the aggregation of a set of individual indicators into a single index, that aims to measure a multi-dimensional concept (OECD, 2008). In the last decades, a growing trend has emerged among policymakers, media, and all areas of research regarding CIs and the rankings derived from them (Bandura, 2011; Hatefi and Torabi, 2010; Greco et al., 2019;

Barbero et al., 2021; Crespo & Crespo, 2016). Examples of CIs encompass diverse domains, such as education, health, government quality, and innovation. The utilization of these ranking systems has yielded substantial ramifications on the perceived performance of decision-making units, including but not limited to countries, regions, cities, and universities (Cherchye et al., 2008; Zabala-Iturriagagoitia et al., 2007a). The influence of these rankings extends beyond mere numerical assessments, as they play a pivotal role in legitimizing the policies adopted by these decision-making units to bolster the development of key sectors such as education, economy, and healthcare. Additionally, they wield the power to shape individuals' inclinations to visit or invest in specific countries.

Besides, the domain of innovation systems holds particular significance, as it has witnessed a growing focus on the development of tools and indicators aimed at evaluating the economic and political capacities of countries. The ultimate objective is to furnish scientific evidence that empowers countries to formulate well-informed policies and undertake strategic initiatives conducive to the advancement, proliferation, and integration of innovation within their societies (Edquist, 2010). The rationale underpinning this emphasis on evaluating innovation capabilities stems from the understanding that fostering a culture of innovation can yield enduring benefits for nations seeking to secure a competitive advantage in the global landscape.

Within this context, CIs assume a key role in providing valuable insights and benchmarks for such endeavors. However, the literature has demonstrated that the methodology employed in constructing CIs significantly influences the final results and rankings (Zabala-Iturriagagoitia et al., 2007a; Grupp and Schubert, 2010; Edquist et al., 2018; Greco et al., 2019; OECD, 2008). Despite this concern, CIs continue to be widely adopted for assessing diverse aspects, including innovation, education, health, and more. The examination of different methodologies to determine whether the results of CIs vary based on data processing, weighting, or aggregation methods is referred to as the robustness of the CI (OECD, 2008).

1.1.2. Innovation Benchmarking

Comparing national innovation performance is a key factor in shaping innovative potential and knowledge capabilities, both of which are integral to the prosperity of a nation. However, it can be difficult for countries to develop and implement innovation strategies individually. Therefore, finding the most appropriate grouping is crucial for any country to help them determine the direction of communication and integration with the relevant group of countries that share similar compositions of innovation enablers and connections. It also helps to enhance the innovation policy, as well as to set goals to fill the gap in comparison to any better-performing group of countries. Moreover, the availability of various grouping perspectives is also significant, as it provides multiple benchmarks and opens new learning opportunities. By determining which countries are more successful at innovation and why nations can maintain or improve their standards by emulating successful behaviors.

The GII offers a comprehensive overview of the current innovation landscape in countries around the world by embracing a detailed data matrix. It groups the countries into four clusters based on the level of income: high-, upper-middle-, lower-middle-, and low-income (Dutta et al.,2019). However, this clustering is not innovation-driven (i.e., it is not based on the GII indicators), but based on the World Bank countries' income insights. As a result, these clusters of the GII do not provide a perceptual map or interpretation of why these countries have been grouped together in terms of innovativeness, nor do they provide technical innovation recommendations.

On the other hand, the European Innovation Scoreboard (EIS) provides four innovation-driven clusters for the European Union (EU) members: leaders in innovation, strongly innovative, moderately innovative, and modest innovators (Guell, 2020). As a result, the EU has implemented several programs such as TEMPUS/TACIS, HORIZON 2020, and ERAZMUS+ to stimulate the integration process regarding socio-economic development at different levels (Liudmyla et al., 2022). The multi-dimensional and innovation-driven clustering approach that is used by the EIS is relevant, as a complex phenomenon such as innovation with latent variables requires a multi-dimensional tool of assessment (Roszko-Wojtowicz and Białek, 2018; Sagan, 2013).

1.2. RESEARCH QUESTIONS AND GOALS

In this matter, out of the multiple CIs that have been developed to assess innovation potential, this study examines the robustness of the 2019 edition of the GII. To do that, we apply alternative data-driven techniques such as Data Envelopment Analysis (DEA), Principal Component Analysis (PCA), Random Forests (RF), and Cluster Analysis (CA) to produce a new rank for the GII countries to examine the robustness of the GII, as well as to produce a new multi-dimensional innovation-driven grouping for the GII countries. Considering the above-mentioned, the following research questions and goals evolved:

Research questions:

1. Does the ranking provided by the GII effectively reflect a country's innovation performance, and can it change with a different perspective?
2. How robust are the results against the methodology used in such CIs?

Research goals:

1. To provide a new evaluation perspective for countries' innovativeness by introducing an efficiency-based ranking for the GII countries.
2. To group the GII countries into different clusters by using a multi-dimensional innovation-driven approach.

3. To reveal the innovation indicators that contribute toward the differentiation of the new clusters.

This dissertation aims to contribute to clarifying these research questions and goals, by focusing on the GII.

1.3. RESEARCH METHOD

To address the research questions and goals outlined earlier, this dissertation adopts a systematic scientific methodology. The methodology comprises two separate streams:

1- The first stream is to construct an efficiency-based ranking for countries that involves utilizing the PCA to generate weights for each indicator. These weights serve as coefficients in a linear combination that produces 21 variables, reflecting the 21 sub-pillars of the GII. Subsequently, the efficiency score for each country is calculated by performing output-oriented DEA seven times, once for each GII pillar. The DEA methodology used in this adjusts for all variables to be treated as output variables and includes a dummy input set to 1, following the Benefit-of-the-Doubt (BOD) approach (Cherchye et al., 2007; Cherchye et al., 2008; Lahouel, 2022; Lafuente, 2022). For instance, the initial efficiency score for a given country is derived by running the DEA with one dummy input equal to 1 and three output variables obtained through the linear combination of the indicators from the “Institutions” pillar. Finally, the overall efficiency score for each country is determined by averaging the results from the seven DEAs conducted for the seven GII pillars. This technique will be referred to as PCA-DEA. Finally, the RF classification method will be utilized to examine the robustness of the PCA-DEA results.

2- The second stream involves clustering the GII countries using a multidimensional innovation-driven approach. This process employs the K-means clustering technique, utilizing the Euclidean distance measurement. To ensure optimal clustering, the Ward.d agglomeration method is employed, minimizing variance within each cluster. Additionally, the Silhouette width analysis is utilized to rectify any incorrect country allocations.

Lastly, the outcomes from both streams will be combined to establish a comparison between the PCA-DEA ranking and the GII ranking. This aggregation of results will enable the extraction of insights and conclusions, facilitating a deeper understanding of the results and their implications. See the methodology model in Figure 1.

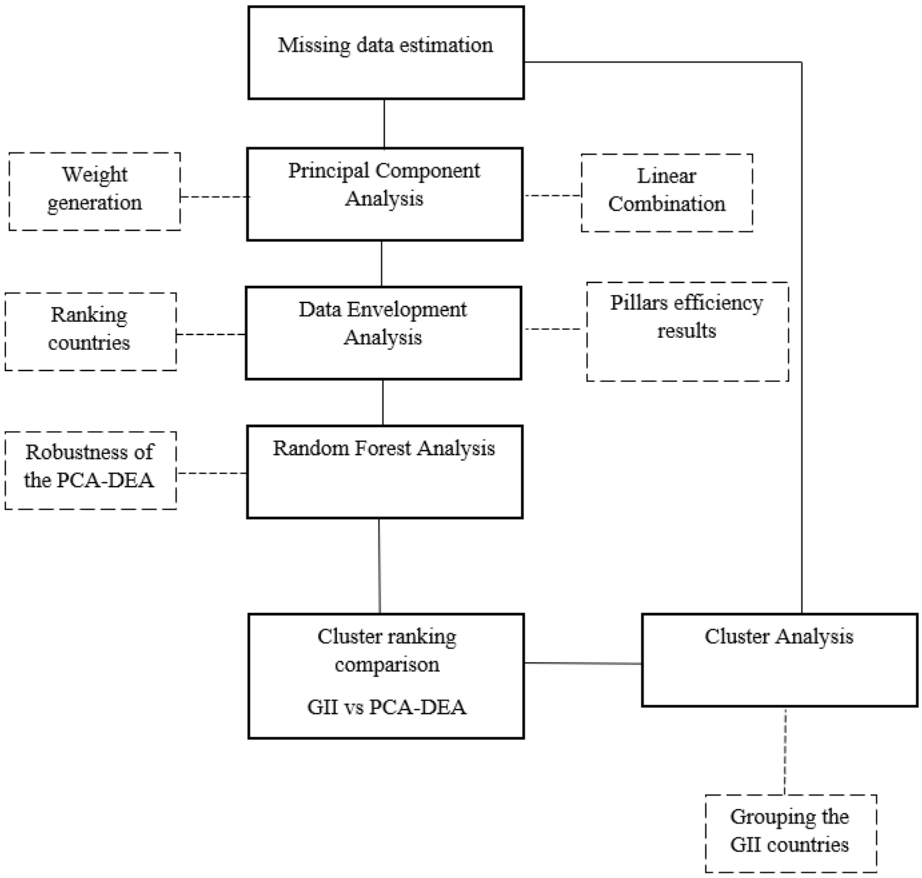

Figure 1. The structure of the Methodology.

Source: own elaboration.

1.4. CONTRIBUTION OF THIS DISSERTATION

This dissertation's main contribution is anchored in the necessity of assessing innovation performance in countries from multiple perspectives, and whether the methodology employed in constructing CIs significantly influences the final results and rankings. Moreover, the finding of the most appropriate clustering is crucial for any country to help determine the direction of communication and integration in the domain of innovation.

Consequently, this dissertation contributes to the existing body of knowledge in several ways. First, the PCA-DEA model introduced in this dissertation is the first efficiency-based ranking for countries' innovation performance at the global level. Second, by using data-driven techniques such as PCA and DEA, the PCA-DEA model introduced in this dissertation was able to produce a robust ranking for the GII countries using intervention-free weights generation. Third, the results of this dissertation confirm the robustness of the GII ranking for both the top and bottom-ranked countries in the GII. However, it evidences that the ranking of countries with medium innovation performance is more sensitive towards the change in the weighting and aggregation schemes.

Fourth, to the best of our knowledge, this dissertation is the first to introduce the RF classification as a tool to examine the robustness of CIs. Fifth, to the best of our knowledge, this dissertation is the first to introduce a multidimensional innovation-driven clustering for countries at the global level. Sixth, the findings of this dissertation suggest that the inefficiency of a resourceful innovation system is related to the size of that system. Consequently, expanding the size of the system might result in more effective resource utilization. Likewise, this principle can be extended to innovation systems with limited resources, prompting them to streamline their innovation efforts toward specific areas of specialization to enhance their efficiency. Finally, the results of this dissertation underscore the potential for reducing the dimensionality of measuring the innovation performance of countries by identifying indicators with minimal impact on capturing structural differences among countries.

1.5. CHAPTER SUMMARY

In this introductory chapter, we have delineated the significance of our research topic and expounded on the rationale behind undertaking an assessment of innovation measurement from diverse perspectives at the national level. Furthermore, we have underscored the imperative of examining various methodologies to ascertain whether CIs yield differing results contingent upon data processing, weighting, or aggregation methods, thus addressing the robustness of the CI.

Moreover, we have highlighted the necessity of introducing a multidimensional innovation-driven clustering approach for countries on a global level. Additionally, we have furnished a brief overview of the research methodology design and the statistical techniques employed in our dissertation. Lastly, we outline the main contributions of this dissertation to the existing body of knowledge.

Chapter 2

THEORETICAL FRAMEWORK

Research is formalized curiosity. It is poking and prying with a purpose- Zora Neale Hurston.

INTRODUCTION

The purpose of this chapter is to provide a theoretical framework that underpins the undertaken doctoral research. The chapter departs in Section 2.1 from the concept of innovation, which is an essential element of economic growth and development (Solow, 1957), introducing its definitions and types. Section 2.2 discusses the concept of national innovation systems, which is an approach that highlights the importance of the interaction between the different actors in the innovation process (Lundvall, 1992). Section 2.3 explores the challenge of measuring innovation due to its latent nature and the importance of innovation-driven clustering for benchmarking and learning among countries. Section 2.4 discusses the GII and highlights the need for developing new perspectives in measuring innovation, given the characteristics underlying the GII. Finally, Section 2.5 discusses the methodological challenges associated with the construction of CIs, highlighting the need for rigorous methods in constructing these, so as to ensure their robustness and reliability. Finally, Section 2.6 summarizes how the chapter provides a comprehensive theoretical grounding guiding the dissertation, highlighting the key concepts and ideas that underpin the dissertation.

2.1. INNOVATION

The portrayal of innovation as a smooth, linear process is a common misconception (Kline, 1985; Rothwell, 1992; Godin, 2006), as it inadequately accounts for the multifaceted and unpredictable causal factors at play (Kline and Rosenberg, 2010). Innovation is a complex and intricate process, subject to various changes and uncertainties (Barbero et al., 2021). Its measurement is also challenging, requiring coordination between technical expertise and market insight, while satisfying multiple economic, technological, and political constraints (Bloch and Bugge, 2013; Gault, 2018), among others. It is crucial to understand that innovation is a holistic process (Borrás and Edquist, 2019), comprising changes not only in "hardware", but also in the market environment, production facilities, and the networks among organizations, knowledge creation (Dutta et al., 2022). Consequently, innovation does not possess a single definition, as the definition of innovation is varying and evolving.

In the early days, innovation was defined by researchers as a new idea, method, or device (Kimberly and Evanisko, 1981; Damanpour and Evan, 1984), and later expanded to include the process of introducing something new (Van de Ven and Rogers, 1988). At a later stage, Kline and Rosenberg (2010) define innovation as not being limited to introducing a new product, 'but it may also be (1) a new process of production (2) the substitution of a cheaper material, newly developed for a given task, in an essentially unaltered product (3) the reorganization of production, internal functions, or distribution arrangements leading to increased efficiency, better support for a given product, or lower costs; or (4) an improvement in instruments or methods of doing innovation' (p. 279). Besides, the 2018 edition of the Oslo Manual (originally introduced in 1992) defines innovation as a 'new or improved product or business process (or a combination thereof) that differs significantly from the firm's previous products or business processes and that has been introduced on the market or brought into use by the firm' (OECD/Eurostat, 2018, p. 20).

The challenge of defining innovation reappears when considering its typology due to the multifaceted nature of the concept. The multiple dimensions and perspectives have resulted in various typologies of innovation. One of the widely acknowledged typologies of innovation is the dichotomy of "Open" versus "Closed", which pertains to the degree of adaptation, ownership, and willingness to engage in external collaboration and resource utilization (Almirall and Casadesus-Masanell, 2010). The “Incremental” versus “Radical” innovation is another typology that is based on the rate of change that occurs within the organization (Herrmann and Peine, 2011). Another typology is the “Technological” versus “Managerial” innovation, which is based on the nature of the activities that are changed (Damanpour and Aravind, 2012). And yet, the list of innovation types continues to business model innovation (Spieth et al., 2014), service innovation (Witell et al., 2016), social innovation (Phillips et al., 2015), etc. Nevertheless, it is important to distinguish between different types of innovation to comprehend the adaptative role that innovation can play in organizations and to develop theories of innovation (Damanpour, 1987; Boschma, 2015).

2.1.1. Open Versus Closed Innovation

The concept of open innovation promotes collaboration among businesses, utilizing external resources and developing dynamism within systems. Conversely, closed innovation emphasizes internal efforts, placing emphasis on ownership and control of proprietary knowledge, as well as independent problem-solving (Chesbrough, 2003). The comparison of open versus closed approaches reveals that an open approach offers a greater potential for a firm to discover new combinations of product features that may be difficult to envision under a closed approach. Nevertheless, when partners have differing objectives, the open approach can restrict the firm’s ability to determine the technological direction of the product (Almirall and Casadesus-Masanell, 2010). Therefore, the optimal innovation approach depends on resolving the trade-off between the advantages of discovery and the costs of divergence.

2.1.2. Incremental versus Radical Innovation

According to Vyas (2005), incremental innovation is a process that is consistently undertaken within an organization to enhance its performance. In the manufacturing sector, this type of innovation is achieved by streamlining and simplifying production through the elimination of waste, with a focus on improving quality and productivity (Cho et al., 2008). Meanwhile, in the service sector, incremental innovation allows for greater employee empowerment and less bureaucracy, which simplifies and enhances customer service (Halliday, 2008). In contrast, radical innovation is characterized by the creation of a new value proposition or market dimension that can provide a significant capability for organizations and develop new strategic operating models that are more strategic than incremental innovation (Christensen and Raynor, 2003). Organizations need to promote both incremental and radical innovation to enhance daily routines as well as a strategic orientation. However, radical innovation often has superiority compared to incremental innovation due to its potential to mobilize the development of new knowledge and technology that solves previously unsolved problems or satisfies customer needs in new ways (Christensen and Raynor, 2003; Oke et al., 2007).

2.1.3. Technological versus Administrative Innovation

Damanpour (1987) contends that it is important to distinguish between technological and administrative (i.e., managerial) innovations, as they pertain to two distinct areas of organizational change, namely technology and social structure. The typology of technological and administrative innovation is significant because it illustrates the differences in the nature of innovations, and also because the two types of innovation together encompass changes that are introduced in a wide range of organizational tasks. Technological innovation plays a significant role in driving technological advancement and enhancing people’s daily lives. It is primarily related to research and development (R&D) activities and is closely linked to the work and production processes within an organization.

Technological innovation pertains to changes in technologies that organizations introduce, such as tools, techniques, physical equipment, or systems, which can help organizations develop their capabilities (Damanpour, 1987). Technological innovation can be categorized as product/service or process technology, as observed in various studies (Yang and Lan, 2010; Ho, 2011; Qian et al., 2013). On the other hand, managerial innovations pertain to changes in administrative processes and organizational structures, which are primarily concerned with management activities, rather than the organization's core activities (Damanpour and Aravind, 2012; Damanpour, 2014).

2.2. NATIONAL INNOVATION SYSTEMS

National Innovation Systems (NIS) have gained increasing significance as a tool for achieving competitive advantage at the country level (Samara et al., 2012). From a Schumpeterian point of view, it is widely acknowledged that the development of a strong innovation capability and efficient performance is a crucial determinant of economic growth. This recognition is particularly salient for economies situated at the middle-income stage, as they are confronted with the challenge of sustaining growth over time (Lavopa and Szirmai, 2018; Lee et al., 2021). The measurement of an NIS's performance is a challenging task, given the latent nature of innovation that cannot be observed directly (Roszko-Wojtowicz and Bialek, 2018). However, various stakeholders, including institutions and scientific communities, have devoted substantial efforts to developing "tools" that can be used to assess the performance of innovation systems. Some of the most notable examples of such tools include the GII (Dutta et al., 2022) and the European Innovation Scoreboard (European Commission, 2023), among others.

NIS refers to the network of institutions, organizations, and actors that interact to promote and support the development and commercialization of new ideas, technologies, and knowledge within a country (Lundvall, 1992; Metcalfe, 1995; Edquist, 1997). This perspective posits that the innovativeness of a country is not solely determined by the individual performance of its constituent organizations and institutions, but also by the ways in which they interact with one another as integrated components of a larger collective system for the creation and utilization of knowledge (Calia et al., 2007; Dahesh et al., 2020). Over the past twenty years, the NIS approach has become the most utilized framework for comprehending the intricate relationships that constitute the innovation process (OECD, 2010). Moreover, empirical studies have emphasized the divergences and functional equivalences that exist in the organizational configurations of NIS across various countries and their implications on the economic performance of each respective nation (Lundvall, 1992; Harding, 2003; Samara et al., 2012).

The systemic approach to innovation is based on the idea that innovation cannot be broken down into isolated phases that occur in a strict linear sequence. Instead, innovation is seen as interactive, collaborative, and interdisciplinary (Huizingh, 2011; Samara et al., 2012). This shift towards a more dynamic and non-linear innovation process is related to the transition from "Mode 1" to "Mode 2" knowledge creation processes (Nowotny et al., 2003), this change in perspective reflects the growing importance of embracing interactive and collaborative approaches to innovation and the rejection of outdated linear models (Limoges et al.,1994). This shift in focus towards NIS has also been accompanied by a transition from a 'closed' to an 'open' innovation approach in the realm of innovation exploration and exploitation (Huizingh, 2011). However, the implementation of a systematic approach to innovation systems does not necessarily entail that such systems can be intentionally designed or strategically planned, given that innovation processes are inherently evolutionary in nature, innovation systems tend to evolve in an unplanned fashion (Edquist, 2010; Dahesh et al., 2020).

In a study conducted by Dahesh et al. (2020), they concluded that contemporary research in the field of innovation systems focuses more on networking, socio-technical transitions, and knowledge exchange, while historical research focused on national, regional (Cooke et al., 1997), sectoral (Breschi and Malerba, 1997), and technological systems (Carlsson and Stankiewicz, 1991). The study suggests several future research directions, such as studying the need to move towards more sustainable systems, exploring strategies for building new trajectories for innovation systems, and investigating the role of social actors and civil society in innovation. In turn, Lee et al. (2021), in their pursuit to classify and analyze the evolution and performance of NIS, focus on economies ranging between middle and high-income. In this regard, they identified five major clusters of NIS that are either balanced or imbalanced in terms of five variables[1]. The study confirms two pathways to achieving catching-up toward high-income status: "balanced catching-up" for economies with balanced and mixed NIS, and "imbalanced and catching-up" for those with imbalanced structures. The study also identifies a “trapped” NIS, consisting of economies perceived to be stuck in the middle-income trap. The imbalanced and catching-up NIS is characterized by the short-cycle time of technologies, low originality, high localization, and high diversification compared with the trapped NIS, which has the opposite attributes. Finally, the balanced and catching-up group is characterized by NIS variables that are balanced at intermediate values.

[1] The five variables are: knowledge localization, technological diversification, cycle time of technologies, originality, and decentralization.

Table 1. Illustration of the five clusters of NIS.

Cluster	Description
Balanced and first tier NIS	Economies with strong and balanced NIS performance, are characterized by high values across all key NIS variables and a low degree of variation.
Balanced and second tier NIS	Economies with a good level of balance in their NIS. While they exhibit balanced attributes across the key NIS variables, their overall performance is not as high as that of the first-tier balanced NIS cluster, and there is some degree of variation among the variables.
Balanced and mixed NIS	Combines economies with moderate values across the five NIS variables, indicating a balanced approach to innovation and development. This cluster includes economies that have achieved successful catching-up and are part of the group of rich economies.
Imbalanced and short cycle NIS	Economies with moderately high overall performance in their national innovation systems. These economies prioritize rapid technological change, knowledge localization, and technological diversification. However, there is notable diversity and imbalance among the five NIS variables in this cluster.
Imbalanced and long cycle NIS	Economies with moderately high overall performance in their NIS. These economies have a slower pace of technological change, lower emphasis on knowledge localization and technological diversification, and significant diversity and imbalance among the five NIS variables.

Reference: Lee et al. (2021).

2.3. MEASUREMENT OF INNOVATION PERFORMANCE

The concept of innovation is complex and ambiguous, with various definitions that are dependent on the nature of socio-economic activities associated with each sector, such as the business, public, or sustainable development sectors. However, the need to adopt a general definition is necessary to develop consistent measurement methods, as well as to introduce new indicators that demonstrate the link and interaction between all actors, and that can be used for evaluating and monitoring the existing innovation policies or to inform new policies (Gault, 2018).

2.3.1. Composite Indicators

In attempting to measure a phenomenon that is not directly observable or measurable, the only way to produce sensible measures is to identify some measurable characteristics that describe this latent variable, referred to as indicators (Panek, 2015; Roszko-Wojtowicz and Bialek, 2018). The existence of these indicators can be used as evidence for the occurrence of the phenomenon under study (Roszko-Wojtowicz and Bialek, 2018). The selection of these indicators, as well as the process of weighting and aggregation to form a composite index, must be supported with an appropriate theoretical framework that justifies the validation and the robustness of the model (Cherchye et al., 2008; OECD, 2018; Greco et al., 2019). The latent nature of the innovativeness (i.e., innovation performance) of a country, prompted scientist communities and various stakeholders to pay efforts to construct composite indices that aim to measure this phenomenon. Some examples of these attempts include the EIS for EU members (at national and regional levels), and internationally the GII at the national level.

The GII is the broadest composite index that aims to measure the innovativeness of countries, covering 129 countries in its 2019 edition (Dutta et al., 2019).[2] It is intended to serve as a benchmarking tool that allows countries to compare their innovation performance with that of their peers and to identify areas where they can learn from each other. The GII involves a detailed data matrix that provides a holistic view of the state of innovativeness for a given country, which can be useful for policymakers and other stakeholders who are looking to understand and promote innovation. This holistic and multi-layering view can allow countries to embrace a comparative analysis of their performance with other relevant peers (Zabala-Iturriagagoitia et al., 2007a; Liudmyla et al., 2022). Additionally, it can facilitate benchmarking countries' innovativeness with others, where a country can gain insights into its relative strengths and weaknesses and identify opportunities for improvement.

However, the final utility of the GII is to provide a final ranking of a set of countries, without any further contribution to policymaking. With this dissertation, we aim to contribute to addressing this particular gap, namely, to allow the GII to become a benchmarking tool and to identify policy areas where countries can intervene through policymaking.

[2] At the moment of writing this dissertation, the latest edition of the GII was the one released in 2022, which included statistical information for 132 countries worldwide. The 2023 edition of the GII is to be released the 27th of September 2023.

2.3.2. Innovation-driven Clustering and Benchmarking

The literature discussed the inadequate use of one single dimension (e.g., income), in clustering units with regard to innovation, besides, the fact that income cannot be regarded as an indicator representing innovation (Jankowska et al., 2017; Stefko et al., 2019; Ma et al., 2022; Liudmyla et al., 2022). To name but a few, Ma et al. (2022) calculated an index for the scientific and technical activities that are necessary for knowledge creation and sustainable growth, hereafter, used all indicators to group all countries into three clusters: leaders, potential leaders, and catching-up countries. Along the same lines, the study by Lee et al. (2021) presented above (see Section 2.2), conducts a cluster analysis of US patent data from 32 to 35 economies to assess and categorize their NIS. As already mentioned (see Table 1), the study identifies five major NIS clusters, some balanced and others imbalanced in terms of the variables used. Furthermore, Jankowska et al. (2017) examine the efficiency of national innovation systems, under the assumption that the higher the GII input, the higher the GII output attained by a country.[3] They divided all countries into three clusters and applied a cross-comparison between the clusters to find that this preassumption was not confirmed. Overall, a complex and latent phenomenon like innovation requires a multi-dimensional clustering that informs learning and policymaking with multiple perspectives.

2.3.3. Dimensionality Reduction

The GII provides a detailed set of data by including 80 indicators for measuring innovation at the national level (Dutta et al., 2019). These large number of indicators reflect the complexity of innovation. However, it has been argued that the GII can be executed by using fewer indicators (i.e., reducing the number of dimensions included in it) (Pence et al., 2018; Cui et al., 2020). Cui et al. (2020) used the Artificial Neural Network and RF to predict

[3] This is similar to the “the more the better” approach for long discussed by Zabala-Iturriagagoitia et al. (2007b), Edquist et al. (2018), or Barbero et al. (2021).

the GII 2019 and 2020 scores by using only 14 indicators out of 80. Subsequently, the possibility of obtaining the GII ranking by using a data-driven approach with a fewer number of indicators would help countries monitor their performance and strategies in a shorter cycle or even in real time. Similarly, Pence et al. (2018) managed to estimate the GII countries' scores by performing the Artificial Neural Network over only 27 indicators of the GII for its 2016 edition. Generally, the widely used techniques for dimensionality reduction are PCA and Factor Analysis, which use a linear transformation of a large dataset to a new coordinate system with fewer dimensions, where most of the variation of the data is captured (Sorzano et al., 2014). Although Cluster Analysis is not commonly used for dimensionality reduction, it can be used by excluding the variables that have negligible mean variation across all clusters.

2.4. THE GLOBAL INNOVATION INDEX AND THE NEED FOR NEW PERSPECTIVES

The GII is a global comparative study that involves 129 countries using detailed evidence of 80 indicators (in its 2019 edition), accounting for around 95% of the global Gross Domestic Product. It provides a tool to measure national innovation performance using two sub-indices, innovation input, and innovation output. There are five pillars under the input sub-index, which consists of Institutions, Human Capital and Research, Infrastructure, Market sophistication, and Business sophistication, while the innovation output sub-index consists of two pillars, Knowledge and Technology Output and Creative Output. Under each pillar, there are several sub-pillars and then under each sub-pillar, there are several indicators [4](Dutta et al., 2019, p. 307), see Figure 2.

The GII represents a significant source of insight in relation to sustainable innovation and aims to evaluate the environment supporting innovation at the national level. At the same

[4] Appendix 1 provides an illustration of the performance shown by Spain in the 2019 edition of the GII, where the previous pillars, sub-pillars and indicators can be observed.

time, it also helps to determine conditions for the diffusion of innovation and its importance for a country's development (Dutta et al., 2019). The GII is distinguished from the other composite innovation indices, as it emphasizes various components related to intangible assets, such as trademarks, business models, and industrial designs (Dutta et al., 2022). Owing to this feature, the GII can be used as a leading reference for researchers, business executives, and policymakers toward creative intangible asset-related innovation.

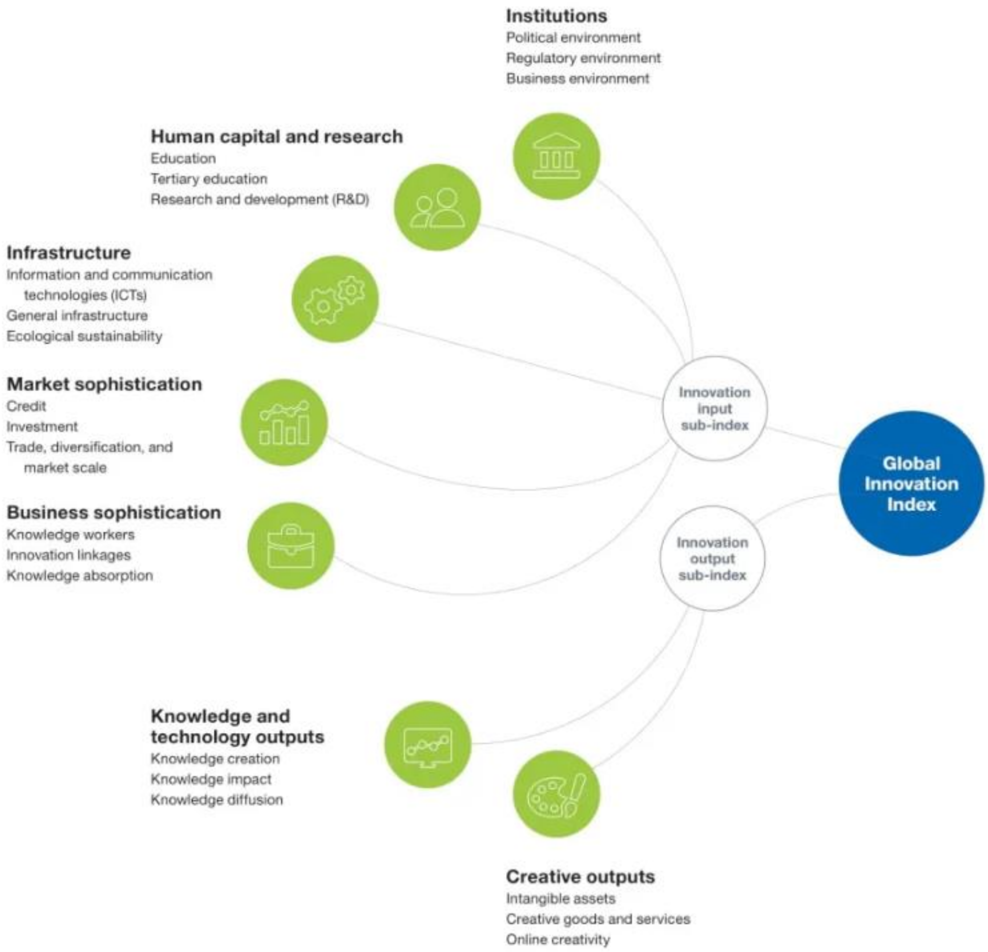

Figure 2. The structure of the GII.

Source: (ForumIAS, 2023)

The GII is not designed to be the ultimate tool to reinforce rankings with respect to economies and countries. Rather, it provides a foundation to be able to continually evaluate the factors of innovation which can provide much insight for economies. In its 2019 edition, it provides a rich database of detailed metrics to redefine innovation policies, which are imperative for policymakers. There are many layers in measuring and understanding innovation, due to its multidimensional nature, which can help identify core and best practices and focus on holistic policies (Zabala-Iturriagagoitia et al., 2007a; Decancq and Lugo, 2013; Edquist et al., 2018). The intricate data allows economies to monitor performance over time and thus standardize developments concerning other economies, allowing comparisons and the identification of best practices, which further supports the aim of the GII. Additionally, the GII 2019 report highlights that it should be "regarded as a sound attempt… to pave the way for better and more informed innovation policies worldwide" (Dutta et al., 2019, p. 387). With this in mind, it is important to recognize that the GII is not the be-all and end-all in rankings within innovation, but a resource that can be used to assist countries in policymaking.

The GII has persistently preserved a similar methodology in all the editions published over the last 10 years, with very slight changes in the weights attributed to the indicators included in it. As a result, it has presented consistency in the results, yielding similar rankings with very small variations across countries.[5] In particular, the aggregation methods followed by the GII are the arithmetic average and weighted average. Also, it attributes the rationality of assigning the weights only to "statistical coherence" and "highest correlation" between indicators, sub-pillars, and pillars (Dutta et al., 2019, pp. 371-373).

In the construction of composite indices, assigning weights is an imperative part of the process, and the level of subjectivity should be treated with caution since it has a bearing consequence on the final results (Greco et al., 2019). In the case of the GII methodology, the allocation and production of weights imply several concerns. First, the GII reveals how weights

[5] This can be confirmed by comparing the results of the 2019 edition of the GII with the 2022 edition (last available year), which yield very similar results (Dutta et al., 2022).

have been generated at the level of the seven pillars, but it has withheld how weights have been generated to its indicators and sub-pillars. Second, the 2019 edition of the GII specifically adjusted the weights for three indicators to provide “statistical coherence”. Third, it needs to be taken into consideration that five of the indicators included in the 2019 edition were collected by subjective means such as using qualitative data (i.e., stakeholder interviews), which might unravel multiple layers of subjectivity (Dutta et al., 2019). Fourth, in the process of data aggregation, the arithmetic average has been used in multiple layers, which is an unreliable statistical tool to employ to combine large amounts of data, since the arithmetic average disregards any significance of certain variables over others.

Interestingly, the GII does not indicate any methodological rationale for using the arithmetic average in the process of aggregation. Furthermore, when aggregating 80 indicators by using averages, there is a high probability that two or more indicators are correlated, creating an additional bias in the results (see Table 2 Section 3.3.2). Retrospectively, duplicating weights may distort the soundness of the outcomes. Considering these methodological concerns, the application of methodologies such as PCA can directly contribute to producing these relative weights in a more statistically objective and sound manner (Adler and Yazhemsky, 2010), as it can help to solve this problem, by grouping indicators along with a set of consistent components, each being given a relative weight.

As previously argued, the GII has presented stability in its rankings for a long period of time, which may not aid in dynamically redefining policies, as the challenges facing innovation systems evolve over time. As things go, it is important to recognize that some countries that follow the direction of the GII, may reap the benefits of its ranking, but this does not necessarily aid the economic growth of that country, as the GII paints all countries with the same brush, disregarding each country on its own merits (Crespo and Crespo, 2016; Jankowska et al., 2017). As a result, the literature has already started to explore new methods for data processing within the GII (Pence et al., 2018; Cui et al., 2020).

The use of unweighted averages between the GII sub-indices to calculate the general index hosts several issues. Firstly, it needs to be considered that the output sub-index disregards the number of innovation inputs utilized. Secondly, it does not seem appropriate to weigh the contribution of the output sub-index equally to the contribution of the input sub-index, since the number of the GII pillars on both sides is not equal. Furthermore, the use of unweighted averages to aggregate the data at the level of pillars/sub-indices, unnecessarily suggests the same significance for all of them within all countries. Accordingly, implying that all countries can be clustered in one big group measured in the same way may yield an unfair scale (Stavbunik and Pelucha, 2019). Additionally, the GII has numerous missing data points, and hence, estimating or imputing these data points may lead to higher precision in results (Cui et al., 2020; Omer et al., 2020).

In determining the GII data structure (input and output), the efficiency of an innovation system can be paired with productivity (i.e., the amount of output a system can generate using a certain amount of input) (Jankowska et al., 2017; Edquist et al., 2018). Hence, in order to compare the efficiency of two systems, it is necessary to measure the output that can be generated by utilizing the same input or less (Charnes et al., 1978; Cherchye, 2007; Barbero et al., 2021; Alnafrah, 2021). In stipulation to this notion, there is always room to improve the GII methodology by using data-driven techniques such as DEA, which seems a very appropriate technique to measure the efficiency of a national innovation system (Hatefi and Torabi, 2010; Omrani et al., 2019; Barbero et al., 2021; Alnafrah, 2021), given that the generation of the weights by using the DEA is taking place without prior intervention, and the efficiency score for each country is relative to all the other countries.

2.5. METHODOLOGICAL CHALLENGES IN THE CONSTRUCTION OF COMPOSITE INDICATORS

CIs play an essential role in formulating the shape of innovation policies and the awareness of them (Edquist et al., 2018). Meanwhile, it is crucial to set the characteristics of the policies that can fulfill the needs of the system (Zabala-Iturriagagoitia et al., 2007a). Therefore, it is alarming to accept and use them simplistically without any scrutiny and discussion (Grupp and Schubert, 2010). Accordingly, there are at least three fundamental challenges that need to be considered when creating CIs: (1) the relative weights associated with the indicators included in them, (2) the methodology used in the aggregation of these indicators, and (3) the robustness of its results (Freudenberg, 2003; Saisana et al., 2005; OECD, 2008; Munda, 2012; Greco et al., 2019). The misuse of the first two factors has a direct impact on the robustness of the final outcomes.

2.5.1. Composite Indicators Weighting

In the process of constructing a CI, selecting or developing the most sensible weighting scheme is critical, due to the strong effect of the weights on the final ranking (Yang et al., 2018; Greco et al., 2019). For example, in the case of the Technological Achievement Index introduced by Desai et al. (2002), changing the weights of some indicators reflected noticeable changes in the overall ranking with consequent (political) implications (Cherchye et al., 2008). According to the OECD (2018), a "weight" represents a coefficient associated with a certain variable in the construction of CIs. In other words, weights represent the contribution of a variable to the overall CI, or to the sub-indices that constitute that overall CI. Therefore, the selection of the weights represents a major challenge in the process of constructing CIs. This challenge is frequently referred to as the 'Index problem' (Cox et al., 1992; Freudenberg, 2003).

To address this challenge a range of weighting schemes have been developed in the literature. For instance, (1) "No or equal weights" suggests that equal weights are assigned for all indicators, which is equivalent to saying that no weights are assigned. This is often labeled in the literature as an Attributes-Based Weighting System (Freudenberg, 2003). This approach has been among others used by the EIS (European Commission, 2019). However, this technique neglects the variation of the relative importance among indicators. Another solution is to adopt (2) Budget Allocation Processes. In this scheme, a set of 'n' points is given to a group of experts so they can distribute them across a group of indicators (Moldan and Billharz, 1997). However, a problem of inconsistency might occur if the number of indicators is larger than 10, or if the group of experts is not carefully selected (Saisana and Tarantola, 2002). This scheme is for example used in the estimation of the weights used in the e-Business Readiness Index introduced by Pennoni et al. (2005). The critical aspect of this scheme is that it involves a high level of subjectivity. Another alternative is to (3) adopt data-driven weight-assigning techniques. In this case, and contrary to the previously mentioned weighting schemes that contain a high level of subjectivity in the arbitrariness of weight selection, statistical techniques such as PCA and DEA are claimed to be desirable as they entail a high level of objectivity in the decision making (OECD, 2008; Decancq and Lugo, 2013).

2.5.2. Aggregation in Composite Indicators

As argued above, characterizing and measuring complex phenomena through a 'simplistic' CI may lead to flawed results and conclusions. An alternative would be to develop several indices that would measure the same phenomenon from different perspectives (Zofio et al., 2023). However, this alternative would increase the complexity of the interpretation of the results, particularly for non-specialized audiences such as the media, or policymakers. They would rather rely on a single number that includes all the indicators, which provides a simple understanding of a complex phenomenon even if this conclusion may be biased and hence lead to wrong (policy) implications (Saltelli, 2007). As a result, the way the CI has been produced is subject to debate. On the other hand, the discussion of the extent to which a CI does genuinely represent the complexity of the phenomenon is also contested (Greco et al., 2019).

An obvious debate emerges in the literature between aggregators versus non-aggregators. Aggregators support building synthetic indices to describe a whole (complex) phenomenon, by combining indicators using a certain aspect to produce a bottom line, which will result in a meaningful outcome (Cobb et al., 1995; Osberg and Sharpe, 2002; Gadrey and Jany, 2003). In turn, non-aggregators consider that the previous aggregation results will be statistically meaningless, and the process must stop at the level of having a set of indicators without combining them to a single outcome, their objection to aggregation being the arbitrariness of the process of weighting and combining (Henderson, 1974; Atkinson et al., 2002). Nevertheless, it is worth stating that most of the widespread indices such as the Human Development Index (UNDP, 2019) adopt a methodological framework that uses aggregation.

Aggregation is the last step in the process of constructing a CI. According to the OECD (2008), aggregation methods can be broken down into three different categories: linear, geometric, and multi-criteria. Moreover, there is yet another categorization of aggregation to be considered, namely, ‘compensatory’ or ‘non-compensatory’ (Munda, 2005). Compensatory approaches occur when there is a trade-off between two perceptions of weights (Paruolo et al., 2013). Understandably, this trade-off might cause a fixed compensability between pairs of dimensions. This could happen when one of the dimensions might cause a loss for another (OECD, 2008). In this context, Munda (2012) argues that in a hypothetical sustainability index, the dimension of economic growth could compensate for the loss in the environmental dimension. To conclude, the Linear or Additive utility-based approach is the most frequently used among the approaches of compensatory aggregation (Saisana and Tarantola, 2002). The non-compensatory multi-criteria approach is less frequently used, due to the simplicity of applying compensatory approaches (Greco et al., 2019). Nevertheless, the Condorcet-Kemeny-Young-Levenglick non-compensatory approach is regarded as a plausible alternative instead of the frequently practiced linear aggregation. This approach has been applied to the Environmental Sustainability Index introduced by Esty et al. (2005), and it has given remarkable differences in the rankings produced by the original compensatory approach (Munda and Nardo, 2005).

2.5.3. Robustness of Composite Indicators

The robustness analysis aims to guarantee the quality and authenticity of a CI during the different levels of its construction, such as the theoretical and methodological framework development (Saisana et al., 2005; Corrente et al., 2021). Therefore, in the absence of robustness analysis, CIs may draw conclusions that would deliver a misleading message to the audience (Saisana et al., 2005; Billaut et al., 2010; Barbero et al., 2021; Corrente et al., 2021). However, the significance of robustness analysis is often neglected by most widespread CIs (OECD, 2008). To ensure a high level of robustness in a CI, some techniques such as uncertainty analysis or sensitivity analysis could be used (OECD, 2008).

Uncertainty analysis refers to the magnitude of changes that may occur in the final outputs of a CI (i.e., the conclusions derived from it) as a result of changes in the construction stages, such as weights assigning or aggregation (OECD, 2008; Grupp and Schubert, 2010; Greco et al., 2019;). In turn, sensitivity analysis deals with how much variance of the final outcomes is yielded due to these uncertainties (Grupp and Mogee, 2004; Saisana et al., 2005). Hereby, it can be concluded that there is a paramount demand to assess the influence of methodological issues amid the construction of CIs. One of the methods that can be beneficial to examine the robustness of CIs, is the classification of the RF. Such classification can be an indicator of the level of a unit (country) ranking.

2.6. CHAPTER SUMMARY

In this chapter, we delved into the definition and significance of innovation for economic growth and development. We explored the various typologies of innovation, such as open versus closed, incremental versus radical, and technological versus managerial. Additionally, we underscored the growing importance of NIS in achieving competitive advantages at the country level. Furthermore, to evaluate and monitor innovation policies effectively, consistent and robust measurement tools are indispensable. However, since innovation is latent and not directly observable, we rely on measurable characteristics, or indicators to do it. Selecting these indicators and the subsequent weighting and aggregation process pose significant measurement challenges.

The GII serves as an imperative tool for measuring innovation at the national level. Covering 129 countries through 80 indicators (in its 2019 edition), the GII offers a comprehensive data matrix providing a holistic view of a country’s innovativeness (i.e., innovation performance). It can serve as a benchmarking instrument, enabling countries to assess their innovation performance relative to peers and identify areas for mutual learning and improvement. Moreover, it can help to identify conditions conducive to innovation diffusion and underscores its significance in a country's development.

While the GII is a valuable resource for policymaking and benchmarking, it is important to recognize that innovation measurement is complex and multifaceted. The GII is just one tool among many, and its methodology has remained consistent over the past decade, using aggregation methods like arithmetic and weighted averages. Nonetheless, there is room for its methodological enhancement, by incorporating data-driven techniques like PCA, DEA, or RF to reduce the inherent subjectivity in indicator weighting and efficiency assessment.

Chapter 3

METHODOLOGY

It is wiser to find out than to suppose – Mark Twain

INTRODUCTION

Methodology plays a vital role in ensuring the credibility and validity of research. It provides a detailed plan that keeps research focused, organized, and on track, resulting in a replicable and effective research process. By outlining the approach and methods used, it allows readers to understand how conclusions are reached. It provides transparency, enabling the evaluation of the study's reliability, and additionally, it highlights any limitations or biases that may influence the findings. In summary, a well-defined research methodology lends legitimacy to the study, generates scientifically sound results, and enhances the reader's comprehension of the process followed during the research.

The structure of this chapter is as follows. Section 3.1 offers a summary of the GII framework and provides specific descriptions of its constituting pillars. Moving on, Section 3.2 discusses the processes involved in the dataset preparation and the estimation of missing data. Lastly, in Section 3.3, we delve into the statistical techniques used in the dissertation, including PCA, DEA, RF, and K-means clustering. Finally, Section 3.4 offers a summary of the contents of the chapter.

3.1. THE GII PILLARS

The GII is structured around seven pillars, namely, Institutions, Human Capital and Research, Infrastructure, Market Sophistication, Business Sophistication, Knowledge and Technology Outputs, and Creative Outputs. These pillars are the fundamental components of the GII, and they collectively provide a holistic and multidimensional assessment of countries' innovation performance[6].

The first five pillars are categorized as input pillars, focusing on the resources and investments for driving innovation. These pillars offer a comprehensive view of a country's readiness and capability for innovation. They represent invaluable information for policymakers, businesses, and researchers in comprehending the fundamental factors that shape a country's innovation potential and its competitive standing on the global stage. The last two pillars of the GII are classified as output pillars, as they represent the outcomes and results of the previous innovation efforts. Also, they provide valuable data to assess the effectiveness of innovation policies, pinpoint areas of excellence, and shape future strategies to enhance a country's innovation landscape. Furthermore, they offer a comprehensive view of how a country's innovation efforts are translated into outcomes and contributions to economic and societal development.

3.1.1. Institutions pillar

Fostering innovation requires having a solid institutional framework. According to Dutta et al. (2019), the indicators included in this pillar are aimed to capture the extent to which countries are attractive for businesses and to stimulate growth through effective governance, adequate protection, and incentivizing measures. Under this pillar, the Political environment sub-pillar

[6] Appendix 1 provides an illustration of the performance shown by Spain in the 2019 edition of the GII, where the previous pillars, sub-pillars and indicators can be observed.

includes two indicators: one focusing on the risks related to political, legal, operational, or security factors affecting businesses, and the other evaluating the quality of public and civil services, policy formulation, and implementation.

The Regulatory environment sub-pillar relies on two additional indicators to gauge the government's capacity in formulating and implementing cohesive policies that facilitate private sector growth. Moreover, it evaluates the extent to which the rule of law prevails in various aspects, such as the cost of redundancy dismissal for workers, contract enforcement, property rights, and judicial effectiveness. Lastly, the Business environment sub-pillar delves into two key factors that directly impact private entrepreneurial activities. It relies on World Bank indices measuring the ease of starting a business and the ease of resolving insolvency. The latter is determined by the recovery rate, represented as the percentage of cents on the dollar recouped by creditors through reorganization, liquidation, or debt enforcement/foreclosure proceedings.

3.1.2. Human capital and research pillar

This pillar aims to gauge innovation potential by assessing the innovation capacity of an economy focusing on its education and research quality. According to Dutta et al. (2019), it gauges the human capital of economies through three sup-pillars. The first sub-pillar measures elementary and secondary education achievements using indicators such as education expenditure and school life expectancy. It also considers the quality of education based on OECD's Programme for International Student Assessment (PISA) results. This pillar emphasizes the progression of higher education, which the second sub-pillar on tertiary education captures through enrollment and prioritizing innovative sectors. The last sub-pillar evaluates research and development (R&D) activities, including the number of researchers, R&D expenditures of top global spenders, and the quality of scientific institutions.

3.1.3. Infrastructure pillar

This pillar assesses the infrastructure that supports eco-friendly communication, transport, and energy and is crucial for fostering the exchange of ideas, services, and goods, leading to enhanced productivity, efficiency, reduced costs, broader market access, and sustainable growth. It consists of three sub-pillars: Information and Communication Technologies (ICTs), General infrastructure, and Ecological sustainability.

The ICT sub-pillar utilizes four indices to assess ICT access, usage, online government services, and citizen online participation. The sub-pillar on general infrastructure uses three indicators: average electricity output per capita (in GWh), a composite indicator measuring logistics performance, and gross capital formation, which includes investments in transportation, education, healthcare, and commercial/industrial buildings. The sub-pillar on ecological sustainability employs three indicators: GDP per unit of energy use, the Environmental performance index, and the number of ISO 14001 certificates issued for environmental management systems compliance.

3.1.4. Market sophistication pillar

This pillar assesses the several factors that influence the prosperity of businesses and the occurrence of innovation, including credit availability, supportive investment environment, access to international markets, competition, and market scale. The pillar is organized into three sub-pillars based on market conditions and overall transaction levels.

The Credit sub-pillar assesses the ease of obtaining credit, considering collateral and bankruptcy laws' role in facilitating lending and the accessibility of credit information. Additionally, it includes data on the total value of domestic credit and the gross loan portfolio of microfinance institutions in emerging markets. The Investment sub-pillar incorporates the ease of protecting minority investors index and two indicators related to transaction levels.

These indicators evaluate whether market size aligns with market dynamism and provide metrics on venture capital deals. The third sub-pillar addresses trade, competition, and market scale. It includes the average tariff rate (weighted by import shares) as an indicator of trade conditions and a survey question reflecting the intensity of competition in local markets (Dutta et al., 2019).

3.1.5. Business sophistication pillar

The enhancement of businesses' productivity, competitiveness, and innovation potential is intrinsically linked to the employment of highly qualified professionals and technicians. This crucial dependence is rooted in the combined factors of market sophistication and skilled human capital. The pillar of business sophistication consists of three sub-pillars.

The knowledge workers sub-pillar includes quantitative indicators such as employment in knowledge-intensive services, availability of formal training at firms, R&D performed by businesses as a percentage of GDP, and the percentage of R&D expenditure financed by businesses. Additionally, it considers the percentage of females employed with advanced degrees, offering insights into gender labour distributions and the sophistication of local human capital.

The innovation linkages sub-pillar evaluates the connections and collaborations between the public, private, and academic sectors that play a vital role in fostering innovation. Also, it involves assessing business/university collaboration on R&D, the existence of well-developed clusters, R&D expenditure funded from abroad, and the number of joint ventures and strategic alliances. Additionally, international linkages are approximated using the number of Patent Cooperation Treaty (PCT) and national office patent applications filed by residents in multiple offices. Measuring innovation linkages poses challenges, prompting the GII team to continually assess various data-based indicators.

The last sub-pillar knowledge absorption is designed to assess how effectively economies absorb and disseminate knowledge. This assessment together with the diffusion of knowledge offer valuable insights into a nation's overall innovation capabilities.

3.1.6. Knowledge and technology outputs pillar

This pillar includes three sub-pillars that encompass variables typically associated with the outcomes of inventions and/or innovations. The first sub-pillar focuses on knowledge creation and includes five indicators resulting from inventive and innovative activities. These indicators comprise patent applications filed by residents at the national and international levels (via PCT), utility model applications at the national level, published scientific and technical articles in peer-reviewed journals, and the number of highly cited articles within an economy.

The second sub-pillar, concerning knowledge impact, presents statistics indicating the effects of innovation activities at both micro- and macro-economic levels, or related proxies. These statistics cover increases in labor productivity, entry density of new firms, spending on computer software, the number of ISO 9001 certificates for quality management systems issued, and the proportion of high- and medium-high-tech industrial output over total manufacturing output.

The third sub-pillar, addressing knowledge diffusion, parallels the knowledge absorption sub-pillar mentioned above. It involves four indicators related to sectors with high-tech content or significant contributions to innovation. These indicators include intellectual property receipts as a percentage of total trade, high-tech net exports as a percentage of total trade, exports of ICT services as a percentage of total trade, and net outflows of foreign direct investment as a percentage of GDP.

3.1.7. Creative outputs pillar

The creative outputs pillar emphasizes the interdependence of creativity and innovation. It consists of three sub-pillars. The first sub-pillar, focusing on intangible assets, incorporates indicators related to valuable brands and trademark applications by residents at the national office, in addition to the industrial designs included in applications at the regional or national level.

The second sub-pillar, addressing creative goods and services, employs proxies to measure the creativity and creative outputs of an economy. Also, it covers heritage and recreational services. Other indicators in this sub-pillar include the production of national feature films, printing and other media output, and creative goods exports, providing insights into the international reach of creative activities in an economy.

The third sub-pillar, focused on online creativity, encompasses four indicators offering deeper insights into the evolution of innovation, production, and trade of digitized creative products and services in innovation-based economies.

3.2. DATASET OF THE STUDY

Due to the COVID-19 pandemic, most countries in the world experience significant changes in the allocation of resources, because of the emergency situation. Hence, and to avoid any potential bias that the COVID-19 pandemic may have caused to countries' innovation indicators, this study uses the data provided by the GII 2019 report (Dutta et al., 2019), which consists of 80 indicators (53 inputs and 27 outputs) collected from 129 countries. The data was available in PDF format only, so we had to transfer it into an MS-Excel sheet manually. It is important to acknowledge that the reliability of the data combines with the element of being presented by well-known international agencies (WIPO, INSEAD, Cornell University) and recognized by institutes and established bodies such as the UN Economic and Social Council as a reliable benchmark for measuring innovation (WIPO, 2021). Furthermore, GII indicators are derived from distinguished socio-economic data such as government effectiveness, tertiary education enrolment, global R&D companies' expenditure, and ICT usage (see Section 3.1).

When analyzing the data of the 2019 GII, there are only sixteen indicators without missing data for all countries. The remaining indicators have different percentages of missing data points, where the highest percentage of missing data for one indicator (High-tech net exports) is 51.2%. Among the 64 indicators with missing data, 27 indicators have less than 5% missing data points, 11 indicators have 5-10% missing data points, 12 indicators have between 10-20% missing data points, 7 indicators have 20-30% missing data points, and the remaining 7 indicators have more than 30% missing data points.

To deal with missing data, a systematic approach was undertaken. Initially, the countries were categorized into four distinct groups based on their income levels, namely High-income, Upper-middle-income, Lower-middle-income, and Low-income, as per the classification provided by the World Bank. Subsequently, the mean value for each indicator was computed within each income group. Additionally, the countries were further classified into four groups based on the Human Development Index (UNDP, 2019). Once again, the mean value for each indicator was calculated within each respective group derived from this classification. Furthermore, to enhance the accuracy of the analysis, each country was grouped together with its five nearest neighboring countries. Consequently, the mean value for each indicator was determined within these geographically proximate groups.

Finally, to obtain an estimation for the missing data points, the arithmetic average of the three aforementioned means was calculated. Nevertheless, despite the comprehensive methodology employed, a total of 40 data points remained missing. To address this issue, a linear regression model was used to impute missing values, thereby ensuring a more complete dataset for subsequent analysis. The model utilized variables that were free of missing data. Within our dataset, a total of 16 variables exhibited complete data for 129 countries, resulting in a sample size of 2064 observations utilized for the imputation process.

3.3. STATISTICAL TECHNIQUES

3.3.1. Data Envelopment Analysis

DEA is a linear nonparametric programming model developed by Charnes et al. (1978) in the field of operations research, which is often referred to as the CCR Model. The idea behind this model is to evaluate the relative efficiency of homogenous Decision-Making Units (DMUs) (i.e., companies, banks, universities, countries, etc.) by giving them scores between 0 and 1. Specifically, DMUs with a score = 1 are considered as "technically efficient" and DMUs with a score < 1 are considered as "inefficient". In the 1980s, Banker et al. (1984) extended the method to deal with contexts requiring multiple inputs and several outputs, which the literature refers to as the BCC-Model. DEA relies on a frontier created from the observed DMUs by utilizing the so-called best-practice, based on the minimum extrapolation principle (Thanassoulis, 2001).

Shen et al.(2013) conclude that DEA provides several features to the field of CIs such as: (1) it is a means to combine multiple indicators for countries without any prior awareness about the tradeoffs (i.e., the weights); (2) the country itself obtains its most favorable relative weights to maximize its efficiency (i.e., minimize inefficiency); (3) if a country is underperforming compared to other countries, this cannot be attributed to the unfair weighting scheme, since every country has been put in its most beneficial position vis a vis all the other countries; and (4) any other weighting scheme would have generated lower weighting scores for that particular country. Additionally, DEA evaluates the relative efficiency of every country, taking into consideration the performance of all other countries (Cherchye et al., 2008). For the above-mentioned features, DEA has been broadly utilized to examine various well-known CIs such as the Technology Achievement Index (Cherchye et al., 2008), the Macro-economic Performance Index (Ramanathan, 2006), the Human Development Index (Despotis, 2005) and the Knowledge Economy composite indicator (Guaita Martínez et al., 2021).

In the context of CI construction, the literature has broadly suggested an adjustment for the classic DEA formulation by considering all the indicators to be treated as outputs (Hermans et al., 2008; Cherchye et al., 2008; Martin et al., 2017; Guaita Martínez et al., 2021). This adjustment is known as the "Benefit-of-the-Doubt" approach (Cherchye et al., 2008; Cherchye et al., 2007; Lahouel, 2022; Lafuente, 2022). It shifts all input variables to become outputs, compromising the inputs with a dummy variable equal to one. It was initially adopted by Melyn and Moesen (1991) as a method to construct CIs to evaluate macroeconomic performance. This approach is to be considered if the underlying structure of the evaluated composite phenomenon is not definitive, or if there is disagreement regarding the construction methodology, or if the input indicators are considered to be "achievements" (Cherchye et al., 2007). All these concerns are valid for any CI that endeavors to measure innovation performance. For example, Crespo and Crespo (2016), by applying a fuzzy-set qualitative comparative analysis, conclude that none of the GII input pillars is a necessary condition for anticipating high innovation performance. Meanwhile, in high-income countries, only two of the pillars (Infrastructure and Human capital and research) are sufficient to secure better innovation performance. Over and above, Jankowska et al. (2017), Edquist et al. (2018), and Barbero et al. (2021) among others, evidence that the common assumption that the higher GII input indicators, the higher GII output indicators is not confirmed (i.e., what Zofio et al. (2023) refer to as the-more-the-better logic). For the above-stated, this study relies on the DEA, using the Benefit-of-the-Doubt approach (see Equations 1 and 2).

$$I^{*}= \quad I^{*}(w)= \quad \max_{I_K,\ k\in\{1,\ldots,M\}} \sum_{q=1}^{Q} I_{qk}\ w_q \qquad \text{Equation (1).}$$

$$CI_c^{*}= \quad \max_{wqc} \sum_{q=1}^{Q} I_{qc}\ w_{qc} \qquad \text{Equation (2).}$$

$S.t.$

$\sum_{q=1}^{Q} I_{qk}\ w_{qk} \ \leq 1$

$w_{qk} \geq 0$

$\forall\, k = 1, \ldots, M\ ;\ \forall\, q = 1, \ldots, Q$

Where I_{qc} is the normalized value of the qth individual variable $(q = 1, \ldots, Q)$ for the country c $(c = 1, \ldots M)$ and w_{qc} the corresponding weight (Cherchye et al., 2004). Whilst I^{*} is the "benchmark performance" (i.e., the hypothetical country that maximizes the overall performance (OECD, 2008)).

Due to the inherent nature of DEA, all efficient DMUs will obtain an identical efficiency score equal to 1 (i.e., DMUs that lay at the frontier). For these DMUs, the final ranking may not be entirely feasible, as they will all achieve the same efficiency score, and hence, it will not be instrumental to make any type of discrimination. This limitation of DEA is known in the literature as the "discrimination power problem" (Adler and Yazhemsky, 2010; Hatefi and Torabi, 2010; Barbero et al., 2021). To address this problem, a sequence of seven sub-DEAs is executed, by dividing the variables for these countries into subsets according to the GII pillars (see Section 3.1). For example, the first sub-DEA will be performed over the output variables: Political environment, Regulatory environment, and Business environment, with a dummy input equal to 1. This will be repeated seven times for the seven pillars. Finally, the total efficiency[7] score for each country will be the average of the seven sub-DEA scores. This approach will also shed light on the pillars individually in terms of the performance of each country within that particular pillar, and the relationship between the individual pillar performance with the overall GII ranking.

3.3.2 Principal Component Analysis

Prior to the application of DEA for each pillar, we conducted a correlation matrix analysis of the indicators within each pillar. This analysis revealed instances of high correlation among various indicators. Table 2 illustrates the correlation matrix for the indicators within the "Institutions" pillar, evidencing high correlations in six instances. This pattern of highly correlated indicators within the same pillar is observed across most of the pillars.

Table 2. The correlation matrix between the indicators that included in the Institutions pillar.

	Political and operational stability	Government effectiveness	Regulatory quality	Rule of law	Cost of redundancy dismissal	Ease of starting a business	Ease of resolving insolvency
Political and operational stability	1.00						
Government effectiveness	**0.88**	1.00					
Regulatory quality	**0.85**	**0.94**	1.00				
Rule of law	**0.87**	**0.95**	**0.94**	1.00			
Cost of redundancy dismissal	-0.18	-0.23	-0.28	-0.25	1.00		
Ease of starting a business	0.37	0.40	0.40	0.39	-0.12	1.00	
Ease of resolving insolvency	0.67	0.72	0.72	0.70	-0.21	0.42	1.00

Source: own elaboration.

Therefore, the necessity for PCA arises to aggregate the indicators within each GII sub-pillar. Ueda and Hoshiai (1997) introduced the PCA-DEA model, emphasizing that PCA was

[7] In the literature, certain researchers opt for the term “Effectiveness” when addressing outcomes associated with the Benefit-of-the-Doubt. This preference is particularly evident in cases where the focus is solely on output settings, as observed in the works of Cherchye et al. (2007) and Jovanović et al. (2022).

introduced to summarize data that include correlated variables while retaining the data's direction of minimum information loss and maximum variance. Consequently, this approach enhances the ability of DEA to discriminate between efficient and inefficient DMUs. Thus, a linear combination is performed to aggregate the indicators under each sub-pillar of the GII into one variable. As a result, the number of indicators will be reduced from 80 to 21, in addition to the 1 dummy input variable.

These 21 output variables are generated by considering the linear combination of the indicators under each sub-pillar of the GII (i.e., $V_q = (u_1.I_1 + u_2.I_2 + \cdots + u_n.I_n)$, I_n= indicators under sub-pillar V_q, $q = (1, \ldots, 21)$, $n =$ number of indicators under each sub-pillar, u_n= the weight generated for indicator n by using the loading of that indicator on the first component[8] of the PCA for the sub-pillar that the indicator belongs to. Eventually, the linear combination of indicators under the sub-pillar Political environment will produce variable 1, and the linear combination of indicators under the sub-pillar Regulatory environment will produce variable 2, etc. (See Table 3). This PCA-DEA approach has been also proposed by Adler and Golany (2002), Adler and Yazhemsky (2010), and Charles et al. (2019).

[8] The first component of a PCA is the component that captures the highest variance in a given sub-pillar. Thus, it carries the most critical information about the sub-pillar.

Table 3. The PCA-DEA output variables.

Sub-pillar	Weight	Variable
Political environment		
Political and operational stability	0.969	**V1**
Government effectiveness	0.969	
Regulatory environment		
Regulatory quality	0.964	**V2**
Rule of law	0.958	
Cost of redundancy dismissal, salary	0.478	
Business environment		
Ease of starting a business	0.842	**V3**
Ease of resolving insolvency	0.842	
Education		
Expenditure on education, % GDP	0.488	
Government funding/pupil, secondary	0.453	
School life expectancy, years	0.894	**V4**
PISA scales in reading, maths, & science	0.881	
Pupil-teacher ratio, secondary	0.761	
Tertiary education		**V5**
Tertiary enrolment, % gross	0.807	

Sub-pillar	Weight	Variable
Trade, competition, & market scale		
Applied tariff rate, weighted avg.	0.716	**V12**
Intensity of local competition	0.809	
Domestic market scale, bn PPP$	0.552	
Knowledge workers		
Knowledge-intensive employment, %	0.914	
Firms offering formal training, % firms	0.332	
GERD performed by business, % GDP	0.817	**V13**
GERD financed by business, %	0.831	
Females employed w/advanced degrees	0.883	
Innovation linkages		
University/industry research collaboration	0.92	
State of cluster development	0.873	
GERD financed by abroad,	0.098	**V14**
JV-strategic alliance deals/bn PPP$ GDP	0.624	
Patent families 2+ offices/bn PPP$ GDP	0.76	
Knowledge absorption		
Intellectual property payments	0.718	**V15**
High-tech imports, % total trade	0.343	
ICT services imports, % total trade	0.514	

Graduates in science & engineering, %	0.725	
Tertiary inbound mobility, %	0.437	
Research & development		
Researchers, FTE/mn pop	0.915	**V6**
Gross expenditure on R&D, %	0.942	
GDP Global R&D companies, avg. exp. top 3	0.927	
QS university ranking, average score top 3	0.881	
(ICTs)		
ICT access	0.914	**V7**
ICT use	0.917	
Government's online service	0.928	
E-participation	0.918	
General infrastructure		
Electricity output, kWh/mn pop	0.849	**V8**
Logistics performance	0.853	
Gross capital formation, % GDP	0.117	
Ecological sustainability		
GDP/unit of energy use	0.646	**V9**
Environmental performance	0.885	
ISO 14001 environmental certificates	0.692	

FDI net inflows, % GDP	0.736	
Research talent, % in business enterprise	0.623	
Knowledge creation		
Patents by origin/bn PPP$ GDP	0.82	**V16**
PCT patents by origin/bn PPP$ GDP	0.871	
Utility models by origin/bn PPP$ GDP	0.325	
Scientific & technical articles/bn PPP$ GDP	0.559	
Citable documents H-index	0.798	
Knowledge impact		
Growth rate of PPP$ GDP/worker	0.074	**V17**
New businesses/th pop. 15-64	0.409	
Computer software spending, % GDP	0.71	
ISO 9001 quality certificates/bn PPP$ GDP	0.691	
High- & medium-high-tech manufactures,	0.801	
Knowledge diffusion		
Intellectual property receipts	0.731	**V18**
High-tech net exports, % total trade	0.377	
ICT services exports, % total trade	0.697	
FDI net outflows, % GDP	0.776	
Intangible assets		**V19**
Trademarks by origin/bn PPP$ GDP	0.496	

Credit		
Ease of getting credit	0.726	
Domestic credit to private sector, % GDP	0.77	**V10**
Microfinance gross loans, % GDP	0.257	
Investment		
Ease of protecting minority investors	0.645	
Market capitalization, % GDP	0.61	**V11**
Venture capital deals/bn PPP$ GDP	0.738	

Industrial designs by origin/bn PPP$ GDP	0.553	
ICTs & business model creation	0.899	
ICTs & organizational model creation	0.881	
Creative goods & services		
Cultural & creative services exports	0.834	
National feature films/mn pop	0.607	
Entertainment & Media market/th pop	0.621	**V20**
Printing & other media, % manufacturing	0.63	
Creative goods exports, % total trade	0.641	
Online creativity		
Generic top-level domains (TLDs)/th pop	0.816	
Country-code TLDs/th pop	0.792	**V21**
Wikipedia edits/mn pop	0.878	
Mobile app creation/bn PPP$ GDP	0.709	

3.3.3. Random Forests

RF is a non-parametric supervised learning statistical method, introduced by Breiman (2001). It has proven to be a reliable method to address classification problems (Hastie et al., 2009; Hamidi and Berrado, 2018). RF develops a random bootstrap of a set of data, performing multiple decision trees according to identified features (variables), eventually by so-called 'Bagging' to vote for the best classification (Hastie et al., 2009). The relationship between CIs and RF emerged recently in the fields of Data mining and Machine learning, to examine the robustness of the classifiers (Setiawan et al., 2019). In this dissertation, the idea behind the use of RF is to ensure the robustness of the PCA-DEA results. By training the RF to classify the countries according to their GII performance, we can then observe the extent to which this classification matches the results of the PCA-DEA.

Practically, all countries have been divided into three groups: (1) Countries in the top quartile of the PCA-DEA ranking are labeled as 'Efficient'; (2) Countries in the lowest quartile of the PCA-DEA ranking are labeled as 'Highly inefficient'; and (3) Countries in the two remaining middle quartiles are labeled as 'Inefficient'. The RF algorithm will be trained by using 15 countries, with 5 countries selected from each group. This training process will enable the RF to capture the characteristics of each group. Subsequently, the RF will be given the task of classifying all remaining countries in each of the previous three groups, based on the patterns and information it has learned. Finally, we will compare the label that we assigned initially to the country with the label that RF has assigned. Another quality RF can provide, is the ability to assess the 'importance' of every variable in the production of the classification (i.e., what are the variables that played an effective role in the classification of the countries?).

3.3.4. Cluster Analysis

To explore the fundamental differences and similarities among the 129 countries included in the GII 2019, K-means cluster Analysis techniques will be utilized. This technique aims to classify the countries into specific clusters based on their innovation indicators (Hollenstein, 2003; OECD, 2008; Ausloos et al., 2018; Bielińska-Dusza and Hamerska, 2021; Liudmyla et al., 2022). The main goal is to divide the 129 countries into homogeneous groups, in which they show a fundamental similarity in the GII indicators. Moreover, it reveals the indicators that contribute the most toward each country's cluster allocation. This allows us to identify indicators that show the innovation structural differences between countries, as well as those that notably influence a country's performance.

Cluster Analysis represents an assortment of statistical techniques that aim to classify certain DMUs (i.e., in our case countries) into clusters, based on the similarity and dissimilarity between them (OECD, 2008). Cluster Analysis emerged initially in the 1930s by Driver and Kroeber (1932). However, it started to gain attention in the 1960s and early 1970s, especially after the introduction of utilizing computers in this field (Blashfield and Aldenderfer, 1978). This technique helped researchers and decision-makers to articulate a meaningful image of the units under study. Particularly, in the context of CIs, one of the main applications of cluster analysis can be found in the Technology Achievement Indicator (Cherchye et al., 2008).

According to Frades and Matthiesen (2010), cluster Analysis can be either hierarchical or non-hierarchical. The hierarchical cluster analysis aims to form a tree dendrogram where the leaves that are seen as most similar by the units are located in the branches that have a shorter distance between them (Frades and Matthiesen, 2010). This dissertation utilizes the K-means non-hierarchical clustering method. This approach is based on the principle that clusters should be formed based on the distances between data points and the centroids of the clusters. Initially, a set of randomly selected centroids is chosen as the starting points for each cluster. The algorithm then iteratively calculates and adjusts the positions of the centroids to optimize the clustering arrangement. K-means clustering is widely utilized in data cluster analysis due to its simplicity of implementation, ability to guarantee convergence, and scalability to handle large datasets (Davidson, 2002; Krishna et al., 2018).

Prior to conducting the analysis, the data underwent a rescaling process using the max-min method. This normalization procedure ensures that variables are on a comparable scale, as all the variables into the 0-1 range. Moving forward in the dissertation, we use the Euclidean distance matrix as shown in Equation (3). Cross-validation tests are then carried out to check the accuracy of clustering results (Kawamoto and Kabashima, 2017). In Table 4 the results of cross-validation tests are reported. As per the K-nearest neighbours algorithm, the connectivity measures to what extent countries in the same cluster are related. It ranges from 0 to infinity and the distance should be kept to the minimum. Also, the measurements of the compactness and separation of the clusters are combined in the Silhouette Width and the Dunn Index. The average of the Silhouette values for each observation is the Silhouette width (Batool and Hennig, 2021). The Silhouette value ranges from [-1,1] and represents the level of confidence in a specific clustering assignment, with a value close to 1, if the data is well-clustered. In turn, the Dunn Index is the ratio between the smallest distance between observations not in the same cluster to the largest intra-cluster distance and it should be maximized. After all, these criteria are to check the internal validity of different clustering techniques (Naghizadeh and Metaxas, 2020). Besides, to assess the stability of the clustering there are other criteria, which include the average distance (AD) between observations within the same cluster, the average distance between cluster centers for observations within the same cluster (ADM), and the average proportion of observations not assigned to the same cluster (APN). The clustering process focuses on the remaining (undeleted) columns, and the figure of merit (FOM) calculates the average intra-cluster variance of the deleted column. In all scenarios, the average of all eliminated columns is utilized, and minimizing all operations is essential.

$$ED_{ij} = \sqrt{\sum_{k=1}^{n}(x_{ik} - x_{jk})^2} \qquad \text{(Equation. 3)}$$

Where:

$ED_{ij} - the\ distance\ between\ the\ country\ i\ and\ j,$

$n - number\ of\ indicators$

$x_{ik} -\ the\ value\ of\ the\ k^{th}\ indicator\ for\ the\ i^{th}\ country,$

$x_{jk} -\ the\ value\ of\ the\ k^{th}\ indicator\ for\ the\ j^{th}\ country,$

Table 4. Clustering cross-validation tests.

Cross-validation criteria	
Internal validation criteria	
Connectivity	52.2361
Dunn index	0.245
Silhouette	0.1915
Stability validation criteria	
The average proportion of non-overlap (APN)	0.1703
The average distance (AD)	1.7558
The average distance between means (ADM)	0.2504
The figure of merit (FOM)	0.1425

Source: own elaboration

Then, we utilize the Ward.d agglomeration method that minimizes the variance within the cluster. According to this method, clusters are grouped according to the smallest distance between clusters (Murtagh and Legendre, 2014). Next, we applied the Elbow method to choose the best clustering number that provides the best validation results. This method consists of two main values, namely, gap statistics and total within the sum square. As shown in Figure 3, the optimal number of clusters is 4. This result mirrors the GII classification framework, where countries are classified into 4 groups[9].

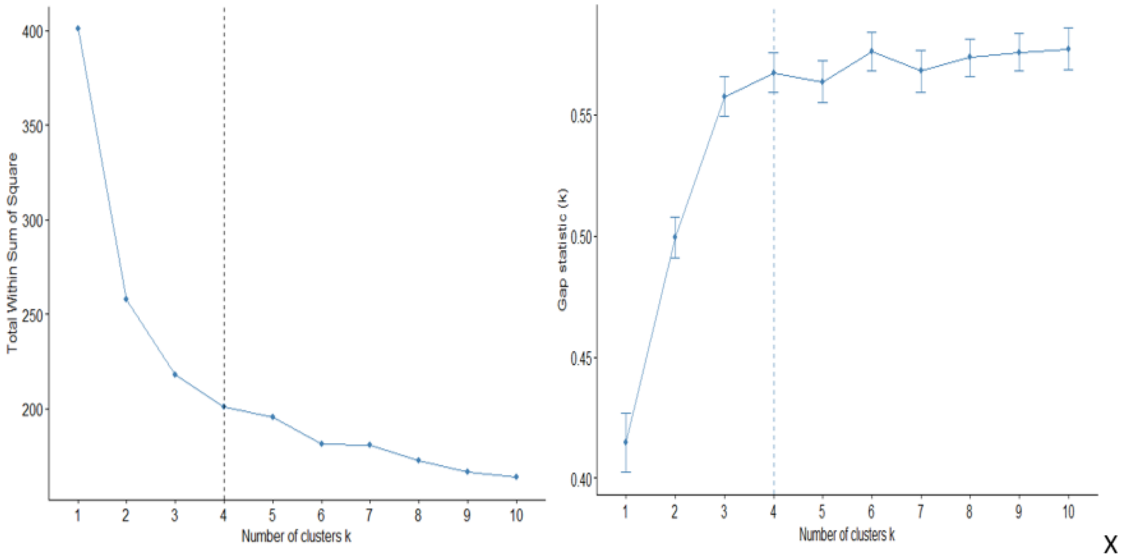

Figure 3. The optimal number of clustering-based Elbow method.
Source: own elaboration.

To identify the GII indicators that contribute the most towards the similarities or dissimilarities between the countries (i.e., the indicators that a particular cluster commonly has a similar level of performance) we decided to find the indicators that make each cluster stand out over the other clusters. This will be done by calculating the mean of every indicator within each cluster and then subtracting it from the overall mean of the variable. Also, by comparing the means of each indicator within each cluster, we can observe the indicators that demonstrate negligible mean variation across all clusters. This analysis enables us to identify the indicators that may have limited discriminative power in distinguishing between different clusters. Moreover, by applying this technique, the variables that contribute most to differentiating each cluster can be identified. Thus, it is possible to identify the innovative strengths and weaknesses

[9] The full list of countries within each cluster can be found in chapter 4.

of each cluster, assisting policymakers in determining the variables on which to focus their innovation policy.

3.4. CHAPTER SUMMARY

In this chapter, we have elucidated the systematic scientific methodology employed in the dissertation to address our research questions and objectives. The chapter begins by emphasizing the significance of data and dataset preparation, including the estimation of missing data. Subsequently, the methodology comprises two separate streams.

Firstly, it involves employing PCA, DEA, and RF to produce a robust efficiency-based ranking of the GII countries. This entails assessing the GII's robustness by comparing the original ranking provided by the GII with the newly derived ranking based on the analysis undertaken in the dissertation. On the one hand, PCA aids in generating weights for each indicator during the aggregation process of each sub-pillar of the GII, a preliminary step before executing DEA. This combination of PCA and DEA is referred to as the PCA-DEA model, which is executed seven times individually, once for each pillar. The final efficiency score for each country is determined as the average of the seven sub-DEAs. Additionally, RF classification is used to examine the robustness of the PCA-DEA results.

Secondly, it entails clustering the GII countries through a multidimensional innovation-driven approach, considering all 80 GII indicators. This is achieved by applying K-means cluster analysis and using Euclidean distance measurement. To ensure optimal clustering, the Ward.d agglomeration method is employed to minimize variance within each cluster. Furthermore, Silhouette width analysis is utilized to rectify any misallocated countries.

Lastly, a comparative analysis of the mean of each indicator within each cluster and a comparison of the mean of each indicator in each cluster with the overall mean of the indicator in the entire dataset are performed. This analysis serves to identify the strengths and weaknesses of each cluster and reveal which indicators play a major role in illustrating innovation structural differences among countries, as well as those notably influencing a country's performance.

Chapter 4
RESULTS

Research means that you don't know, but are willing to find out - Charles F. Kettering

INTRODUCTION

This chapter presents the key findings derived from the research methodology. In what follows, Section 4.1 presents notable findings emanating from the PCA, which was utilized to generate the weights for the indicators included in the GII. Subsequently, Section 4.2 expounds upon outcomes stemming from the PCA-DEA model, harnessed for the computation of efficiency scores and the ensuing efficiency-based rankings for all countries. To determine the robustness of the PCA-DEA model, Section 4.3 conducts an RF analysis. The multidimensional approach to clustering GII countries is elucidated in Section 4.4, where the K-means clustering technique is used. This not only facilitates the categorization of countries into different clusters but also unveils the strengths and weaknesses inherent in each delineated cluster. Finally, Section 4.5 summarizes the main findings of the empirical contribution of the dissertation.

4.1. PRINCIPAL COMPONENT ANALYSIS

As discussed in Section 3, methods such as PCA are useful to generate the weights for the linear combination V_q. Thus, a PCA was performed for every GII sub-pillar separately. The loading of every indicator on the first component of the PCA is the weight of that indicator in the linear combination V_q. As a result, the GII 80 indicators were aggregated into 21 variables, one variable for each sub-pillar (i.e., 3 variables for each pillar - see Appendix 2).

Table 3 in Section 3 provides the value of weights associated with each indicator. The majority of the indicators received weights exceeding 0.5, while certain indicators like "GERD financed by abroad" and "Growth rate of PPP$ GDP/worker" received notably low weights, falling below 0.1. Sub-pillars from the input side of the GII demonstrated strong representation in the linear combination, with an average weight of approximately 0.73. Particularly noteworthy were sub-pillars like "Political environment," "Research & development," and "ICTs," which received weights exceeding 0.9. In contrast, sub-pillars from the output side had an average weight of 0.67, slightly lower than their input side.

The results show that indicators in the GII sub-pillar "Political environment" have gained an equal weight of 0.97. This case of nearly equal weights among all the indicators in a given sub-pillar has occurred in several sub-pillars such as "Business Environment", "Research & Development", "Information & Communication Technologies", "Investment", and "Online creativity".

4.2. DATA ENVELOPMENT ANALYSIS

After the application of the linear combination V_q outlined earlier to aggregate the 80 indicators of the GII into 21 variables, the Benefit-of-Doubt DEA technique was employed seven times, one for each pillar (i.e., Institutions, Human capital and research, Infrastructure, Market sophistication, Business sophistication, Knowledge and technology outputs, and Creative outputs). Finally, the overall PCA-DEA efficiency score for each country was determined by averaging the results of the seven DEAs, as illustrated in Table 5 The closer the score to 1[10], the higher the level of efficiency in the national innovation system of the respective country.

[10] The initial outcomes generated by the Benefit-of-Doubt method are expressed as ratios equal to or greater than 1, with 1 representing the best performance and scores >1 indicating lower efficiency. To facilitate the reading, we have modified the results by reciprocating the ratios to align them with the well-known DEA format, which typically ranges between 0 and 1, where 1 represents the highest efficiency.

The table employs a color-coded scheme, where green indicates a high-efficiency score and red indicates a low-efficiency score. This visual representation facilitates the tracking of efficiency levels across the countries and pillars simultaneously. Furthermore, Table 5 reveals that most of the countries demonstrate higher efficiency scores in the input pillars of the GII as compared to the output pillars. Notably, the pillar "Human Capital and Research" exhibits the highest average efficiency score of 0.80 among all countries. Conversely, the pillar "Market Sophistication", in comparison to the input pillars of the GII, exhibits the lowest average efficiency score of 0.44. On the output side of the GII, the pillar "Knowledge and Technology Outputs" displays a lower efficiency score with an average of 0.36, representing the lowest score among all the pillars.

In the context of the "Human Capital and Research" pillar, which has the highest-efficiency average observed across all countries suggests that inefficient countries are relatively close to the efficiency frontier. This implies that the indicators within this pillar may not effectively capture performance variations among countries. A quick examination of the 12 indicators within this pillar reveals, that indicators like "School life expectancy," "Tertiary enrollment," "Graduates in science & engineering," and "Tertiary inbound mobility" have limited growth potential with no big room for a large variation.

Conversely, within the "Knowledge and Technology Outputs" pillar, where the average efficiency is the lowest, there is a notable gap between countries on the efficiency frontier and those falling within it. This suggests that the indicators within this pillar are capable of capturing variations in performance. Notably, these indicators, such as "Patents by origin," "Scientific & technical articles," "High- & medium-high-tech manufactures," and "ICT services exports, % total trade," are sophisticated and advanced, contributing to the observed performance differentiation.

Table 5. The results of the seven PCA-DEAs.

	Country	Institutions	Human Capital and Research	Infrastructure	Market Sophistication	Business Sophistication	Knowledge and Technology Outputs	Creative Outputs	Overall PCA-DEA efficiency
1	Albania	0.8585	0.7612	0.7352	0.3612	0.4514	0.2173	0.4334	**0.5455**
2	Algeria	0.6874	0.6862	0.6056	0.1300	0.4094	0.0970	0.3561	**0.4245**
3	Argentina	0.6652	0.8694	0.7300	0.2429	0.4722	0.2759	0.5454	**0.5430**
4	Armenia	0.7570	0.7615	0.6707	0.4182	0.5025	0.2238	0.6781	**0.5731**
5	Australia	0.9641	0.9942	0.9593	0.8068	0.7509	0.6889	0.7857	**0.8500**
6	Austria	0.9024	0.9239	0.9090	0.4833	0.8499	0.4964	0.7701	**0.7621**
7	Azerbaijan	0.8634	0.7642	0.7183	0.3488	0.6700	0.1186	0.5294	**0.5732**
8	Bahrain	0.7246	0.8610	0.8668	0.4142	0.6513	0.2263	0.4988	**0.6061**
9	Bangladesh	0.5886	0.7122	0.5681	0.2581	0.4550	0.0992	0.3740	**0.4365**
10	Belarus	0.7883	0.8789	0.8325	0.2804	0.6493	0.5411	0.5497	**0.6458**
11	Belgium	0.9552	0.9289	0.8580	0.4551	0.8827	0.5272	0.7582	**0.7665**
12	Benin	0.7090	0.6704	0.3905	0.1858	0.4034	0.0919	0.3443	**0.3993**
13	Bolivia	0.5756	0.7713	0.5914	0.3804	0.3549	0.1122	0.3652	**0.4501**
14	Bosnia and Herzegovina	0.6879	0.7744	0.5400	0.4127	0.4032	0.6093	0.3802	**0.5440**
15	Botswana	0.7137	0.7920	0.5768	0.2996	0.4365	0.2858	0.3724	**0.4967**
16	Brazil	0.6949	0.7468	0.8333	0.3806	0.5662	0.3141	0.5376	**0.5819**
17	Brunei Darussalam	0.8546	0.8789	0.7627	0.4776	0.5931	0.1517	0.4660	**0.5978**
18	Bulgaria	0.7716	0.8139	0.8286	0.3983	0.5901	0.9517	0.6887	**0.7204**
19	Burkina Faso	0.6971	0.6419	0.4635	0.2167	0.3985	0.0908	0.3438	**0.4075**
20	Burundi	0.6771	0.6130	0.2982	0.2004	0.4492	0.0554	0.2780	**0.3673**
21	Cambodia	0.5577	0.7300	0.4566	0.5841	0.5670	0.1148	0.4964	**0.5010**
22	Cameroon	0.6636	0.6982	0.4394	0.2538	0.4592	0.0984	0.3630	**0.4251**
23	Canada	0.9864	0.9557	0.9294	0.6505	0.7833	0.6598	0.8162	**0.8259**
24	Chile	0.8278	0.8383	0.8135	0.5858	0.5383	0.3815	0.6477	**0.6618**
25	China	0.8062	0.9376	0.7953	0.8984	0.8488	1.0000	1.0000	**0.8981**

26	Colombia	0.8245	0.7797	0.7944	0.4941	0.5309	0.4218	0.5073	**0.6218**
27	Costa Rica	0.6410	0.7797	0.7755	0.5043	0.5777	0.1999	0.6970	**0.5964**
28	Côte d'Ivoire	0.7651	0.7535	0.4692	0.3320	0.3447	0.1143	0.4445	**0.4605**
29	Croatia	0.7553	0.8669	0.7934	0.3895	0.5458	0.6105	0.5770	**0.6483**
30	Cyprus	0.9001	0.8244	0.8841	0.9073	0.7265	1.0000	0.7575	**0.8571**
31	Czech Republic	0.8839	0.9061	0.8664	0.4193	0.6832	0.7651	0.7115	**0.7479**
32	Denmark	0.9767	1.0000	1.0000	0.8212	0.8802	0.6767	0.8573	**0.8874**
33	Dominican Republic	0.6528	0.6490	0.7125	0.2548	0.4810	0.0853	0.5073	**0.4775**
34	Ecuador	0.5184	0.7711	0.6554	0.2683	0.4343	0.2054	0.4912	**0.4777**
35	Egypt	0.6825	0.7378	0.6528	0.3199	0.5078	0.1752	0.4592	**0.5050**
36	El Salvador	0.6695	0.7757	0.5988	0.4517	0.3511	0.1380	0.5426	**0.5039**
37	Estonia	0.8736	0.9524	0.9257	0.4718	0.6760	0.8602	0.9289	**0.8126**
38	Ethiopia	0.5470	0.6007	0.4457	0.3554	0.5297	0.0874	0.3321	**0.4140**
39	Finland	1.0000	0.9949	0.9510	0.5542	0.9157	0.6635	0.8935	**0.8532**
40	France	0.9039	0.9164	0.9722	0.5299	0.8007	0.6963	0.8841	**0.8148**
41	Georgia	0.8386	0.7460	0.6886	0.5091	0.4610	0.2453	0.5593	**0.5783**
42	Germany	0.9581	0.9352	0.9502	1.0000	0.9378	0.8356	0.8840	**0.9287**
43	Ghana	0.6006	0.7003	0.5955	0.2523	0.6309	0.0959	0.4182	**0.4705**
44	Greece	0.7981	1.0000	0.8622	0.5237	0.5324	0.6114	0.5424	**0.6957**
45	Guatemala	0.6172	0.7019	0.5612	0.3865	0.5478	0.0770	0.5445	**0.4909**
46	Guinea	0.6641	0.5615	0.4867	0.1918	0.7337	0.0485	0.4949	**0.4545**
47	Honduras	0.5896	0.7207	0.5314	0.4883	0.5217	0.1465	0.5292	**0.5039**
48	Hong Kong, China	0.9858	0.9729	0.9459	1.0000	0.9818	0.6720	0.8347	**0.9133**
49	Hungary	0.7768	0.8698	0.7851	0.3702	0.7649	0.6143	0.6263	**0.6868**
50	Iceland	0.9417	0.8948	1.0000	0.5215	0.7512	0.4891	1.0000	**0.7998**
51	India	0.6577	0.6523	0.6659	0.4939	0.7446	0.3815	0.4940	**0.5843**
52	Indonesia	0.8051	0.7302	0.5800	0.3732	0.6969	0.1301	0.5243	**0.5486**
53	Iran	0.5583	0.8050	0.6460	0.4028	0.4747	0.2734	0.8289	**0.5699**
54	Ireland	0.9449	0.9322	1.0000	0.3938	1.0000	0.8863	0.7843	**0.8488**
55	Israel	0.8915	1.0000	0.8839	0.4539	0.9430	0.8401	0.9495	**0.8517**
56	Italy	0.9006	0.8907	0.9154	0.4412	0.7691	1.0000	0.6488	**0.7951**

57	Jamaica	0.9028	0.7564	0.6352	0.3092	0.5849	0.1075	0.6779	**0.5677**
58	Japan	0.9671	0.9643	0.9703	0.7814	0.9417	0.9130	0.8483	**0.9123**
59	Jordan	0.6225	0.7376	0.6473	0.3857	0.6004	0.1802	0.5287	**0.5289**
60	Kazakhstan	0.8683	0.7647	0.8117	0.3259	0.5173	0.1038	0.4551	**0.5495**
61	Kenya	0.7549	0.7188	0.4948	0.4115	0.7052	0.1829	0.5510	**0.5456**
62	Kuwait	0.6517	0.8134	0.8087	0.4726	0.5487	0.1926	0.5142	**0.5717**
63	Kyrgyzstan	0.7592	0.7571	0.5992	0.3350	0.3500	0.0955	0.3143	**0.4586**
64	Latvia	0.8299	0.8938	0.7878	0.4997	0.5414	0.5399	0.7476	**0.6914**
65	Lebanon	0.5842	0.6782	0.6422	0.5082	0.5445	0.5758	0.3816	**0.5592**
66	Lithuania	0.7984	0.8666	0.8352	0.3838	0.6281	0.3962	0.7592	**0.6668**
67	Luxembourg	0.9595	0.8792	0.9728	0.4285	0.9001	1.0000	1.0000	**0.8771**
68	Macedonia	0.6026	0.6836	0.7395	0.3740	0.4217	0.4058	0.5089	**0.5337**
69	Madagascar	0.6604	0.6102	0.3615	0.1865	0.4642	0.0922	0.4613	**0.4052**
70	Malawi	0.5967	0.6701	0.5005	0.3423	0.4035	0.1014	0.3054	**0.4171**
71	Malaysia	0.8171	0.7672	0.8477	0.6749	0.8744	0.4080	0.6031	**0.7132**
72	Mali	0.6890	0.6276	0.4578	0.2012	0.5587	0.0892	0.3587	**0.4260**
73	Malta	0.7945	0.8566	0.9676	0.4012	0.7553	0.5038	0.9122	**0.7416**
74	Mauritius	0.8823	0.7787	0.7131	0.5825	0.5328	0.3202	0.5059	**0.6165**
75	Mexico	0.8461	0.7744	0.7805	0.4284	0.5939	0.2525	0.5659	**0.6060**
76	Mongolia	0.6346	0.8011	0.6197	0.4787	0.3484	0.1408	0.9130	**0.5623**
77	Montenegro	0.8245	0.7748	0.7351	0.4600	0.5181	0.2275	0.6139	**0.5934**
78	Morocco	0.7873	0.7836	0.7108	0.3764	0.4690	0.1706	0.5584	**0.5509**
79	Mozambique	0.6183	0.6495	0.4537	0.1813	0.4759	0.1000	0.3890	**0.4097**
80	Namibia	0.6359	0.7136	0.6288	0.4274	0.4958	0.0769	0.6491	**0.5182**
81	Nepal	0.7106	0.6982	0.5708	0.4558	0.4257	0.1393	0.3965	**0.4853**
82	Netherlands	1.0000	0.9404	0.9752	0.5498	1.0000	1.0000	0.9235	**0.9127**
83	New Zealand	1.0000	0.9362	0.9650	0.8772	0.7372	0.4880	0.8546	**0.8369**
84	Nicaragua	0.6528	0.7315	0.5709	0.2571	0.3811	0.0964	0.4264	**0.4452**
85	Niger	0.7187	0.6241	0.3839	0.1193	0.4644	0.0881	0.3411	**0.3914**
86	Nigeria	0.6123	0.6992	0.5579	0.3368	0.4093	0.0885	0.4375	**0.4488**
87	Norway	0.9975	0.9362	1.0000	0.7034	0.8119	0.4633	0.7967	**0.8156**

88	Oman	0.7300	0.8925	0.8073	0.3846	0.6785	0.1359	0.5467	**0.5965**
89	Pakistan	0.7657	0.7047	0.4354	0.2143	0.5725	0.1484	0.4497	**0.4701**
90	Panama	0.6040	0.7980	0.7226	0.5770	0.4996	0.1396	0.6057	**0.5638**
91	Paraguay	0.6415	0.7796	0.5972	0.2802	0.3776	0.1160	0.8098	**0.5145**
92	Peru	0.6917	0.7319	0.7308	0.4092	0.4388	0.1583	0.5131	**0.5248**
93	Philippines	0.6868	0.7941	0.7360	0.1899	0.7046	0.3424	0.5478	**0.5717**
94	Poland	0.8607	0.9219	0.8723	0.4381	0.6022	0.4370	0.5620	**0.6706**
95	Portugal	0.9228	0.9156	0.9098	0.5204	0.6731	0.6151	0.8167	**0.7676**
96	Qatar	0.7122	0.7399	0.8245	0.4100	0.7952	0.1361	0.5555	**0.5962**
97	Republic of Korea	0.9654	1.0000	1.0000	0.7322	0.9756	0.9943	0.8755	**0.9347**
98	Republic of Moldova	0.8083	0.7885	0.7703	0.3328	0.3574	0.1876	0.6856	**0.5615**
99	Romania	0.7765	0.8030	0.7874	0.3620	0.4855	0.6805	0.5161	**0.6301**
100	Russian Federation	0.8186	0.8991	0.8611	0.4555	0.6306	0.3306	0.5682	**0.6520**
101	Rwanda	0.8024	0.7985	0.5232	0.4033	0.5510	0.0774	0.4349	**0.5129**
102	Saudi Arabia	0.5369	0.8713	0.7746	0.3457	0.6755	0.1638	0.5168	**0.5549**
103	Senegal	0.7246	0.5803	0.5214	0.2069	0.4937	0.1106	0.4813	**0.4455**
104	Serbia	0.8283	0.8104	0.7845	0.3707	0.5095	0.6456	0.5160	**0.6378**
105	Singapore	1.0000	1.0000	0.9591	0.7104	0.9284	0.5831	0.7022	**0.8405**
106	Slovakia	0.8079	0.8490	0.8330	0.4469	0.5641	0.5939	0.6791	**0.6820**
107	Slovenia	0.9536	0.9353	0.8328	0.3115	0.9871	0.6163	0.8343	**0.7816**
108	South Africa	0.7327	0.7873	0.7056	0.7310	0.6843	0.3546	0.4858	**0.6402**
109	Spain	0.8963	0.9032	0.9534	0.5770	0.6268	0.6632	0.7314	**0.7645**
110	Sri Lanka	0.7181	0.7358	0.7114	0.2982	0.5183	0.1361	0.4444	**0.5089**
111	Sweden	0.9800	0.9609	0.9950	0.6539	1.0000	0.7257	0.9536	**0.8956**
112	Switzerland	0.9841	0.9289	1.0000	0.8258	1.0000	0.9343	0.9884	**0.9516**
113	Tajikistan	0.6566	0.6323	0.5073	0.2800	0.5299	0.0595	0.3655	**0.4330**
114	Tanzania	0.6031	0.6143	0.5124	0.2736	0.6032	0.0930	0.4280	**0.4468**
115	Thailand	0.9141	0.7874	0.6501	0.7453	0.6615	0.2974	0.5368	**0.6561**
116	Togo	0.7311	0.6613	0.4525	0.2464	0.4642	0.1315	0.3665	**0.4362**
117	Trinidad and Tobago	0.7403	0.7882	0.7018	0.3633	0.4520	0.1356	0.4586	**0.5200**
118	Tunisia	0.7802	0.7295	0.7204	0.4738	0.4614	0.2786	0.4767	**0.5601**

119	Turkey	0.6960	0.8626	0.7837	0.4997	0.5694	0.2585	0.6418	**0.6160**
120	Uganda	0.6058	0.7042	0.4666	0.2611	0.5575	0.1049	0.3768	**0.4396**
121	Ukraine	0.6631	0.7982	0.6207	0.3880	0.5451	0.2812	0.7219	**0.5740**
122	United Arab Emirates	0.8210	0.8667	0.9536	0.5154	0.9255	0.2197	0.5807	**0.6975**
123	United Kingdom	0.9444	0.9326	0.9978	0.7350	0.8922	0.9173	0.8869	**0.9009**
124	USA	0.9833	0.9027	0.9670	1.0000	1.0000	0.8437	0.8467	**0.9348**
125	Uruguay	0.7711	0.7993	0.8753	0.2963	0.4441	0.3961	0.6258	**0.6011**
126	Viet Nam	0.6722	0.9240	0.6165	0.7192	0.5172	0.3164	0.6202	**0.6265**
127	Yemen	0.5016	0.6916	0.5668	0.0200	0.3131	0.0323	0.2528	**0.3397**
128	Zambia	0.5913	0.7031	0.5157	0.3582	0.4489	0.0755	0.3409	**0.4334**
129	Zimbabwe	0.4957	0.6248	0.4350	0.3088	0.3405	0.1739	0.2867	**0.3808**

Source: own elaboration.

Table 5 also reveals that the infrastructure pillar possesses the highest number of efficient countries, totaling 6. In contrast, the Market Sophistication and Creative Outputs pillars have the lowest number of efficient countries, each with only 3 nations. On a different note, the Netherlands stands out as an efficient performer in 3 pillars: Institutions, Business Sophistication, and Knowledge and Technology Outputs. Several other countries, including but not limited to Switzerland, the USA, the Republic of Korea, and Denmark, demonstrate technical efficiency in 2 pillars. Additionally, China and Luxembourg were the only two efficient countries in the 2 pillars of the output side of the GII. On the other hand, Iceland is the sole country efficient in 2 pillars, with one on the input side and one on the output side of the GII. Overall, Switzerland, Sweden, and the Netherlands secure positions in the top ten across more than 5 pillars, whereas countries like Bulgaria and Slovenia make a single appearance in the top ten of any of the pillars.

Furthermore, some other countries have the potential to significantly enhance their innovation performance by focusing on improvement in just one pillar. For instance, Australia, Austria, Canada, and Denmark could remarkably boost their efficiency scores by concentrating on enhancing their performance in the Knowledge and Technology Outputs pillar. Similarly, the Netherlands, Ireland, Israel, and Luxembourg could achieve substantial improvements in their performance by prioritizing the Market Sophistication pillar. In the case of countries like Qatar, Trinidad and Tobago, and Jordan, any progress in the output pillars of the GII (i.e., Knowledge and Technology Outputs and Creative Outputs), could lead to a significant enhancement in their innovation performance.

Lastly, Table 6 below demonstrates a significant correlation between the GII ranking and the countries' efficiency scores in each pillar. It also shows that the highest correlation is observed between the GII ranking and the infrastructure pillar, while the lowest correlation is found with the Market Sophistication pillar.

Table 6. Correlation between the GII ranking and each pillar's efficiency score.

GII pillars efficiency score	GII Ranking		
	Correlation	R square	Significance F
Institutions	0.79	0.62	0
Human Capital and Research	0.87	0.76	0
Infrastructure	0.91	0.82	0
Market Sophistication	0.73	0.53	0
Business Sophistication	0.78	0.60	0
Knowledge and Technology Outputs	0.84	0.71	0
Creative Outputs	0.87	0.75	0

Source: own elaboration.

Based on the obtained results, we can proceed to establish a comprehensive ranking encompassing all 129 countries in the GII. To facilitate comparison, Table 7 presents the PCA-DEA ranking of each country compared to the GII 2019 ranking. Notably, the sign of the difference holds a meaning, where a negative value indicates moving backward in the new ranking, while a positive value signifies moving forward. Furthermore, within the table, two columns (the GII score and the difference) are color-coded. In the case of the GII score, green signifies a higher ranking, while red indicates a lower ranking. Similarly, for the difference column, green denotes progression in the PCA-DEA rank, whereas red signifies regression in the PCA-DEA rank. The comparison is undertaken from two perspectives: (1) by analyzing the difference between a country's position in the GII rank and its position in the PCA-DEA rank, and (2) by examining the distribution of these absolute differences throughout the GII rank.

Table 7. Comparison between the final PCA-DEA rank and GII 2019 rank for all the countries.

Country	PCA-DEA	GII	Diff.	Abs.Value	Country	PCA-DEA	GII	Diff.	Abs.Value
Switzerland	1	1	0	0	Azerbaijan	66	84	18	18
United States of America	2	3	1	1	Armenia	67	64	-3	3
Republic of Korea	3	11	8	8	Kuwait	68	60	-8	8
Germany	4	9	5	5	Philippines	69	54	-15	15
Hong Kong, China	5	13	8	8	Iran	70	61	-9	9
Netherlands	6	4	-2	2	Jamaica	71	81	10	10
Japan	7	15	8	8	Panama	72	75	3	3
United Kingdom	8	5	-3	3	Mongolia	73	53	-20	20
China	9	14	5	5	Republic of Moldova	74	58	-16	16
Sweden	10	2	-8	8	Tunisia	75	70	-5	5
Denmark	11	7	-4	4	Lebanon	76	88	12	12
Luxembourg	12	18	6	6	Saudi Arabia	77	68	-9	9
Cyprus	13	28	15	15	Morocco	78	74	-4	4
Finland	14	6	-8	8	Kazakhstan	79	79	0	0
Israel	15	10	-5	5	Indonesia	80	85	5	5
Australia	16	22	6	6	Kenya	81	77	-4	4
Ireland	17	12	-5	5	Albania	82	83	1	1
Singapore	18	8	-10	10	Bosnia and Herzegovina	83	76	-7	7
New Zealand	19	25	6	6	Argentina	84	73	-11	11
Canada	20	17	-3	3	Macedonia	85	59	-26	26
Norway	21	19	-2	2	Jordan	86	86	0	0
France	22	16	-6	6	Peru	87	69	-18	18
Estonia	23	24	1	1	Trinidad and Tobago	88	91	3	3
Iceland	24	20	-4	4	Namibia	89	101	12	12
Italy	25	30	5	5	Paraguay	90	95	5	5
Slovenia	26	31	5	5	Rwanda	91	94	3	3
Portugal	27	32	5	5	Sri Lanka	92	89	-3	3
Belgium	28	23	-5	5	Egypt	93	92	-1	1
Spain	29	29	0	0	Honduras	94	104	10	10
Austria	30	21	-9	9	El Salvador	95	108	13	13
Czech Republic	31	26	-5	5	Cambodia	96	98	2	2
Malta	32	27	-5	5	Botswana	97	93	-4	4
Bulgaria	33	40	7	7	Guatemala	98	107	9	9
Malaysia	34	35	1	1	Nepal	99	109	10	10
United Arab Emirates	35	36	1	1	Ecuador	100	99	-1	1
Greece	36	41	5	5	Dominican Republic	101	87	-14	14
Latvia	37	34	-3	3	Ghana	102	106	4	4
Hungary	38	33	-5	5	Pakistan	103	105	2	2
Slovakia	39	37	-2	2	Cote d'Ivoire	104	103	-1	1
Poland	40	39	-1	1	Kyrgyzstan	105	90	-15	15
Lithuania	41	38	-3	3	Guinea	106	125	19	19
Chile	42	51	9	9	Bolivia	107	110	3	3
Thailand	43	43	0	0	Nigeria	108	114	6	6

Russian Federation	44	46	2	2	Tanzania	109	97	-12	12
Croatia	45	44	-1	1	Senegal	110	96	-14	14
Belarus	46	72	26	26	Nicaragua	111	120	9	9
South Africa	47	63	16	16	Uganda	112	102	-10	10
Serbia	48	57	9	9	Bangladesh	113	116	3	3
Romania	49	50	1	1	Togo	114	126	12	12
Viet Nam	50	42	-8	8	Zambia	115	124	9	9
Colombia	51	67	16	16	Tajikistan	116	100	-16	16
Mauritius	52	82	30	30	Mali	117	112	-5	5
Turkey	53	49	-4	4	Cameroon	118	115	-3	3
Bahrain	54	78	24	24	Algeria	119	113	-6	6
Mexico	55	56	1	1	Malawi	120	118	-2	2
Uruguay	56	62	6	6	Ethiopia	121	111	-10	10
Brunei Darussalam	57	71	14	14	Mozambique	122	119	-3	3
Oman	58	80	22	22	Burkina Faso	123	117	-6	6
Costa Rica	59	55	-4	4	Madagascar	124	121	-3	3
Qatar	60	65	5	5	Benin	125	123	-2	2
Montenegro	61	45	-16	16	Niger	126	127	1	1
India	62	52	-10	10	Zimbabwe	127	122	-5	5
Brazil	63	66	3	3	Burundi	128	128	0	0
Georgia	64	48	-16	16	Yemen	129	129	0	0
Ukraine	65	47	-18	18					

Source: own elaboration.

The comparison between the GII 2019 and PCA-DEA rankings reveals a discernible disparity in their respective rankings. Specifically, the average absolute difference between the two ranks amounts to 7.2 positions. Notably, certain countries undergo substantial shifts in their positions, such as Mauritius and Macedonia, experiencing changes of more than 20 ranks. Conversely, countries such as Switzerland and Yemen maintain their relative standings without any alterations.

Furthermore, the results demonstrate that countries positioned in the middle of the GII ranking exhibit the most significant discrepancies in their absolute values of difference between the two rankings. These countries demonstrate substantial changes in their rankings, both advancing and regressing. Conversely, countries placed at the extremities of the GII ranking, belonging to either the top quartile or lower quartile, demonstrate minimal shifts in their rankings, with none of them experiencing changes exceeding 10 positions (see Figure 4). This could be explained by the Attributes-based weighting scheme (i.e., no weights or equal weights) that the GII uses. It is worth noting that what may appear important for one group of countries may not hold the same importance for others, as emphasized by studies like Zabala-Iturriagagoitia et al. (2007a) and Stavbunik and Pelucha (2019). Also, the GII uses the unweighted average of the GII sub-indices to calculate the general index. Instead, the PCA-DEA relies on the objective weights provided by the PCA and DEA to aggregate the GII indicators and sub-pillars.

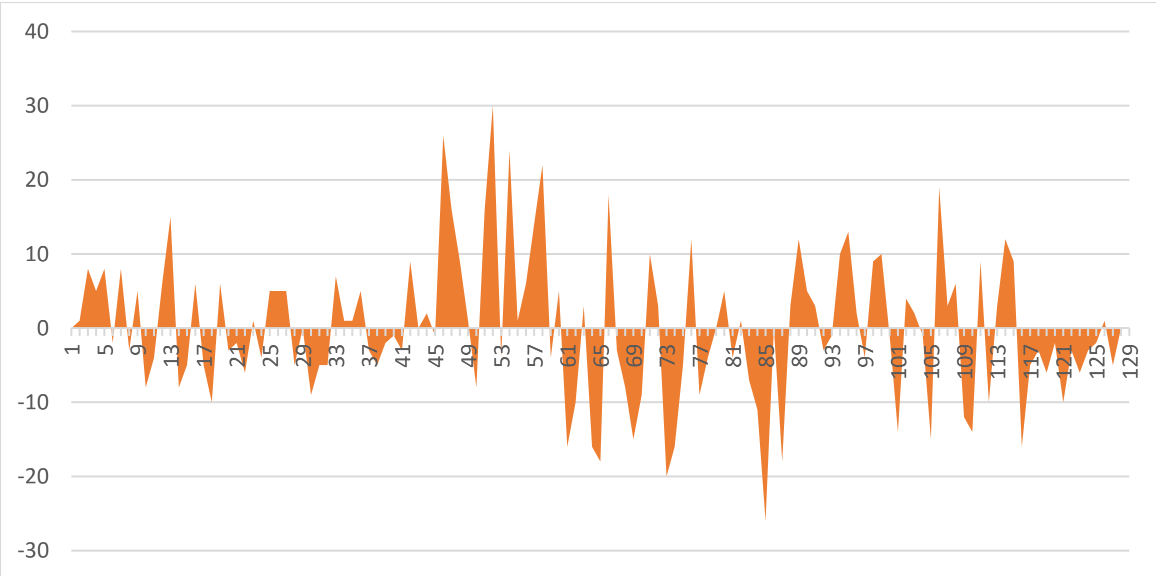

Figure 4. The difference between the GII Ranking and the PCA-DEA ranking.

4.3. RANDOM FORESTS

To ensure the robustness of the PCA-DEA results, the RF technique was applied. To elaborate, all countries were divided into three groups. The first group consists of the top 32 countries in the PCA-DEA rank, this group being labeled as "Efficient. The second group consists of the lowest 32 countries in the PCA-DEA rank, this group is labeled as "Highly inefficient". Lastly, the third group consists of the remaining 65 countries in the middle of the PCA-DEA rank, which is labeled as "Inefficient".

The application of the RF analysis yielded a congruence of 87.5% between the PCA-DEA ranking and the RF classification. To elaborate, among the top 32 countries in the PCA-DEA ranking that have been labeled as "efficient", the RF also classified 30 of them as efficient, indicating a high degree of agreement. The RF classification of the two remaining countries (Cyprus and Malta) differed, reclassifying them as inefficient. This can be referred to their moderate performance in the indicator "Researchers FTE/mn pop" and their very low performance in the indicator "QS university ranking", which are the most important 2 indicators in the RF classification, as presented in Figure 5.

Regarding the highly inefficient countries, out of the lowest 32 countries in the PCA-DEA ranking, the RF classified 23 of them as highly inefficient. Nevertheless, there were 9 countries that the RF reclassified as inefficient. These 9 countries are Algeria, Bolivia, Cote d'Ivoire, Dominican Republic, Ecuador, Ghana, Guatemala, Kyrgyzstan, and Tajikistan. The reasons behind this upgrade are multifaceted. However, a unifying thread among these countries is that they all have better performance than the others in the indicators "Electricity output" and "Government's online service".

In the middle of the PCA-DEA ranking, encompassing 65 countries that were labeled "Inefficient", the RF classified 60 of them also as inefficient. However, the RF classified the remaining 5 countries as highly efficient. These 5 countries are Cambodia, Colombia, Honduras, Namibia, and Rwanda. A salient commonality among these countries is that they all have a notable low performance in indicators that are considered important for the RF classification, such as "Researchers FTE/mn pop" and "Logistics performance". This high level of agreement of the RF indicates the robustness of the PCA-DEA model to produce an efficiency innovation index for the GII countries. An overview of these results is presented in Table 8.

Table 8. Random forests classification results.

PCA-DEA / RF	Efficient	Inefficient	Highly Inefficient	Class. Error
Efficient	30	2	0	0.06
Inefficient	0	60	5	0.08
Highly Inefficient	0	9	23	0.28
	OOB estimate of error rate= 12.4%			

Source: own elaboration.

Moreover, the application of RF analysis not only contributes to classifying countries according to their innovation performance but also sheds light on the relative importance of individual indicators in determining the classification outcomes. Figure 5 illustrates the importance of each indicator, evidencing the extent of influence exerted by each indicator in the classification process. For instance, the indicators "Researchers, FTE/mn pop", "ICT access" and "QS university ranking" emerge as the most influential indicators, significantly impacting the classification results. Notably, the figure further reveals that the 10 most important indicators, in terms of their impact, are all derived from the input side of the GII. On the output side of the GII, indicators such as "Entertainment & Media market", "ICTs & organizational model creation", and "ISO 9001 quality certificates" hold paramount importance in determining the classification outcomes. This observation underscores the essential role played by these particular indicators in assessing and characterizing the innovation output of the GII countries. On the other hand, the strong agreement observed between the RF and PCA-DEA underscores the RF's effectiveness in forecasting a country's innovation performance. This predictive capacity could potentially entail the utilization of the important indicators, as delineated in Figure 5 by the RF. For instance, the most important 10 indicators serve as a viable proxy for gauging and comparing countries' levels of innovation performance.

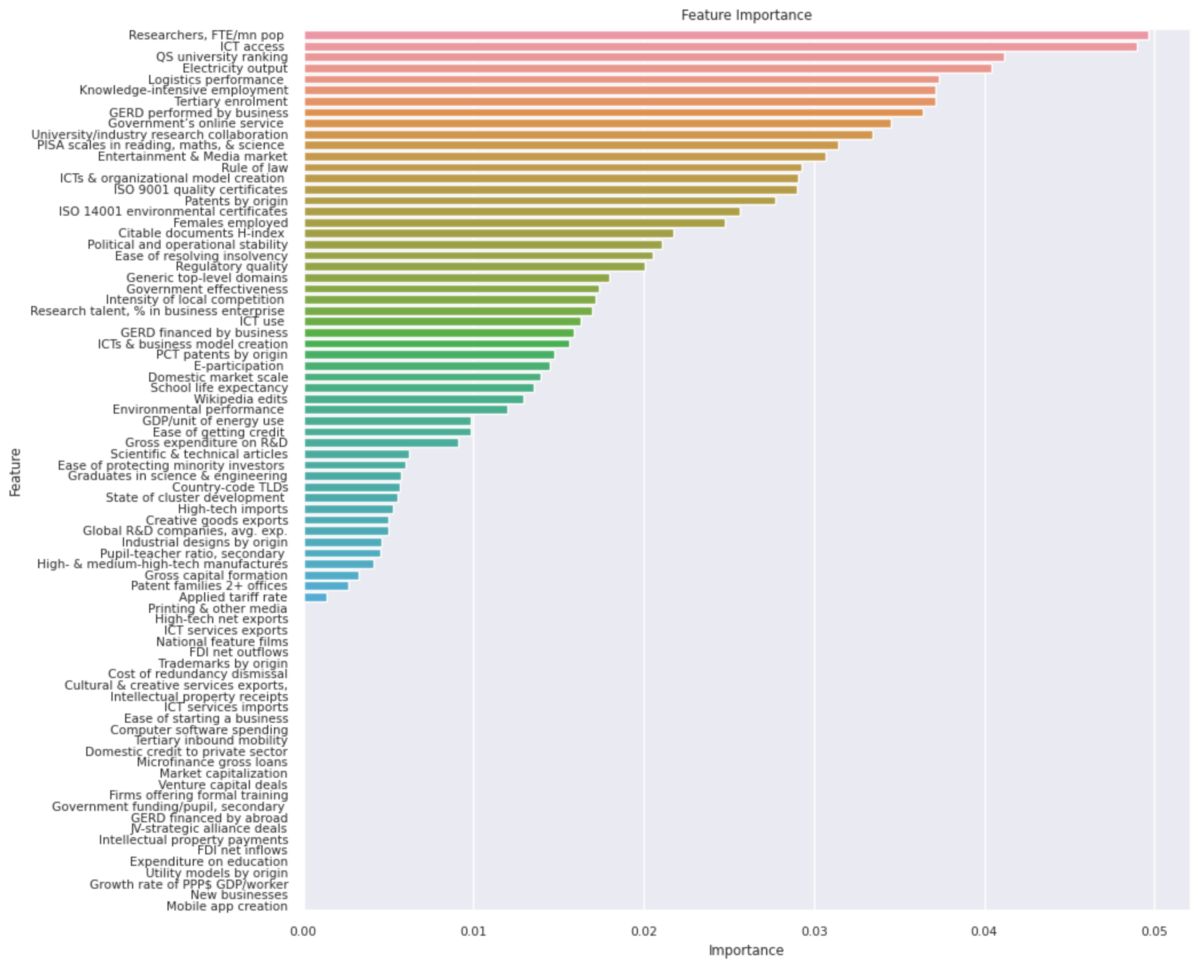

Figure 5. The level of importance for every GII indicator in determining the RF classification[11].

4.4. CLUSTER ANALYSIS

Finally, a K-means cluster analysis was applied in accordance with the methodology outlined in Section 3.3.4, resulting in the distinctive initial grouping of countries into four clusters, as illustrated in Figure 6. Specifically, the following clusters are obtained: Cluster 1 comprises 37 countries, Cluster 2 encompasses 50 countries, Cluster 3 consists of 19 countries, and Cluster 4 includes 23 countries. Based on the characteristics and descriptive statistics of these clusters, we have assigned labels to each of them (see Appendix 3). Cluster 1 has been

[11] The indicator importance in the RF is determined by observing the change in error when permuting the values of an indicator. If permuting the values causes a considerable change in the error, it signifies the importance of the indicators within the model.

labeled as the "Primitives", Cluster 2 as the "Intermediates", Cluster 3 as the "Advanced", and Cluster 4 as the "Specials". Refer to Table 9 for further details.

Table 9. The initial 4 clusters that obtained by applying the K-means clustering technique.

The Primitives	The Intermediates	The Advanced	The Specials
Algeria	Albania	Bulgaria	Australia
Bangladesh	Argentina	China	Austria
Benin	Armenia	Croatia	Belgium
Bolivia	Azerbaijan	Cyprus	Canada
Botswana	Bahrain	Czech Republic	Denmark
Burkina Faso	Belarus	Estonia	Finland
Burundi	Bosnia and Herzegovina	Greece	France
Cambodia	Brazil	Hungary	Germany
Cameroon	Brunei Darussalam	Italy	Hong Kong, China
Cote d'Ivoire	Chile	Latvia	Iceland
El Salvador	Colombia	Lithuania	Ireland
Ethiopia	Costa Rica	Malaysia	Israel
Ghana	Dominican Republic	Malta	Japan
Guatemala	Ecuador	Poland	Luxembourg
Guinea	Egypt	Portugal	Netherlands
Honduras	Georgia	Slovakia	New Zealand
Kenya	India	Slovenia	Norway
Madagascar	Indonesia	Spain	Republic of Korea
Malawi	Iran	United Arab Emirates	Singapore
Mali	Jamaica		Sweden
Mozambique	Jordan		Switzerland
Namibia	Kazakhstan		United Kingdom
Nepal	Kuwait		United States of America
Nicaragua	Kyrgyzstan		
Niger	Lebanon		
Nigeria	Mauritius		
Pakistan	Mexico		
Paraguay	Mongolia		
Rwanda	Montenegro		
Senegal	Morocco		
Tajikistan	North Macedonia		
Togo	Oman		
Uganda	Panama		
United Republic of Tanzania	Peru		
Yemen	Philippines		
Zambia	Qatar		
Zimbabwe	Republic of Moldova		
	Romania		
	Russian Federation		
	Saudi Arabia		

Serbia
South Africa
Sri Lanka
Thailand
Trinidad and Tobago
Tunisia
Turkey
Ukraine
Uruguay
Viet Nam

Source: own elaboration.

However, by visualizing the clustering results plot in Figure 6, it is apparent that there is an overlap between the first and second clusters, reflecting a clustering error. To detect the countries that are placed in the wrong cluster and causing those overlaps, a Silhouette width analysis is conducted (Batool and Hennig, 2021). The results of the Silhouette width analysis of the K-means clustering are presented in Figure 7.

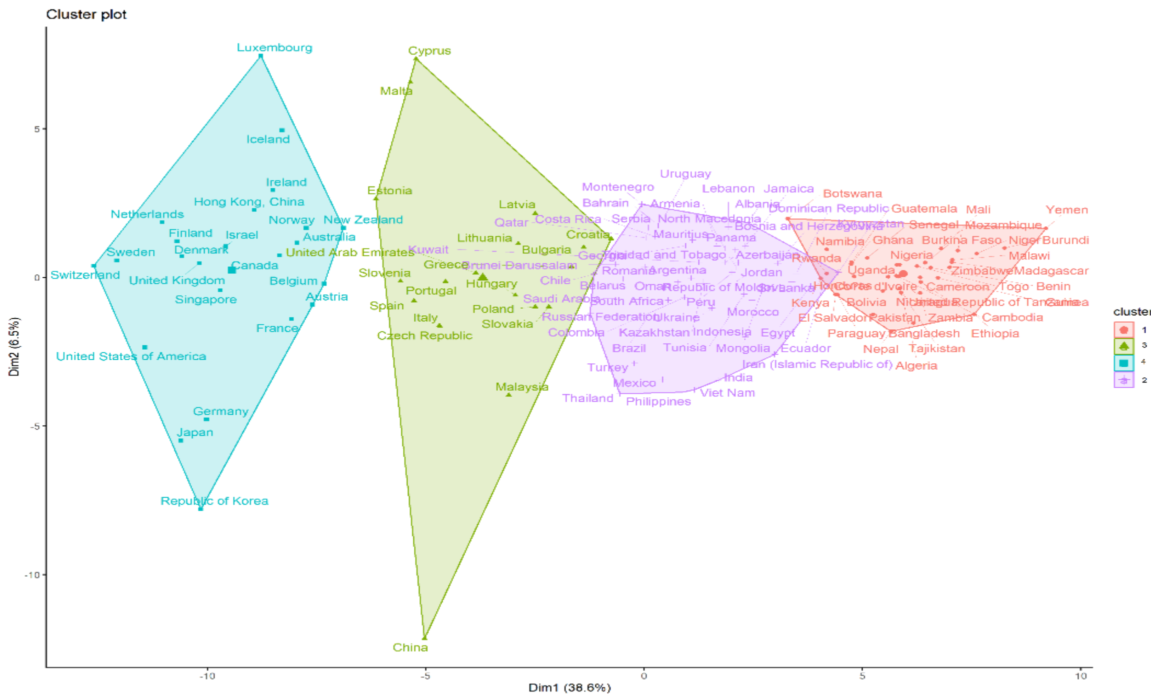

Figure 6. K-means Clustering results plot[12].

[12] The two dimensions of this plot are the first two principal components to define the X-Y coordinates of each observation.

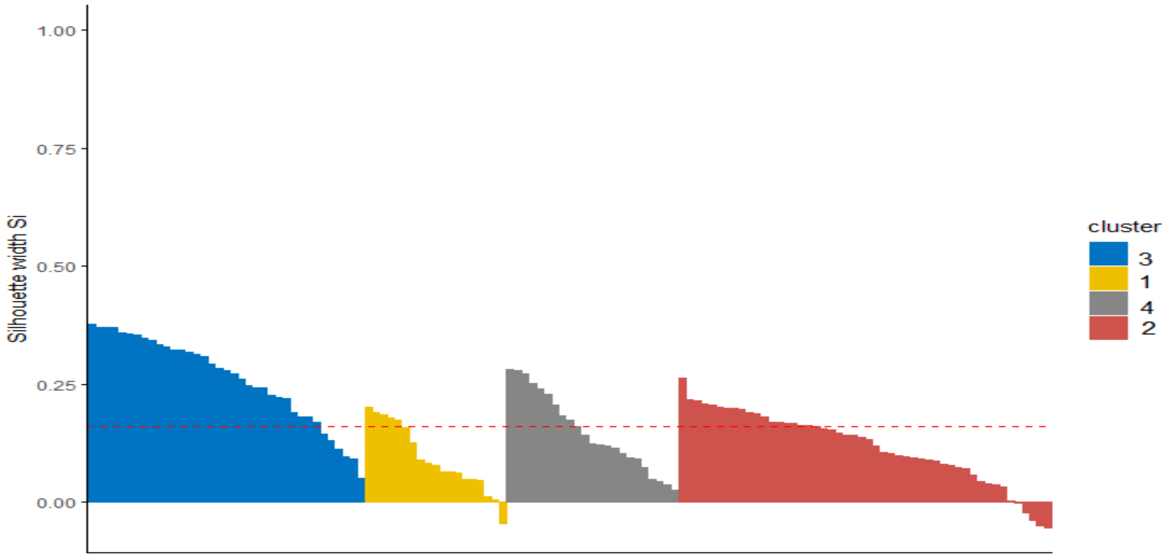

Figure 7. Silhouette width of K-means clustering.

The results of the Silhouette width of K-means clustering confirm the existence of overlapping mainly in the first and second clusters. Table 10 shows the list of the 6 countries that are subject to overlapping, where the "Cluster" column denotes the cluster to which each country is currently assigned, while the "Neighbour" column signifies the cluster to which the country should ideally belong. Particularly, the results indicate that countries such as Egypt and Ecuador are positioned as intermediates within the clusters. However, when considering the Silhouette width error, these countries are ultimately classified as Primitives in relation to their overall innovation performance.

Table 10. List of wrongly clustered countries in the K-means clustering.

Country	**Cluster**	**Neighbour**	**Silhouette width**
Croatia	3	2	-0.0438
Sri Lanka	2	1	-0.00177
Jamaica	2	1	-0.02108
Kyrgyzstan	2	1	-0.03846
Egypt	2	1	-0.05022
Ecuador	2	1	-0.05283

Source: own elaboration.

The final step at this stage is to make corrections by changing the location of the 6 countries that are wrongly clustered to their appropriate clusters, as indicated in Table 10. See Table 11 for the final corrected results of K-means clustering.

Table 11. The final clustering after the corrections suggested by Silhouette width analysis.

The Primitives	**The Intermediates**	**The Advanced**	**The Specials**
Algeria	Albania	Bulgaria	Australia
Bangladesh	Argentina	China	Austria
Benin	Armenia	Cyprus	Belgium
Bolivia	Azerbaijan	Czech Republic	Canada
Botswana	Bahrain	Estonia	Denmark
Burkina Faso	Belarus	Greece	Finland
Burundi	Bosnia and Herzegovina	Hungary	France
Cambodia	Brazil	Italy	Germany
Cameroon	Brunei Darussalam	Latvia	Hong Kong, China
Cote d'Ivoire	Chile	Lithuania	Iceland
Ecuador	Colombia	Malaysia	Ireland
Egypt	Costa Rica	Malta	Israel
El Salvador	Croatia	Poland	Japan
Ethiopia	Dominican Republic	Portugal	Luxembourg
Ghana	Georgia	Slovakia	Netherlands
Guatemala	India	Slovenia	New Zealand
Guinea	Indonesia	Spain	Norway
Honduras	Iran (Islamic Republic of)	United Arab Emirates	Republic of Korea
Jamaica	Jordan		Singapore
Kenya	Kazakhstan		Sweden
Kyrgyzstan	Kuwait		Switzerland
Madagascar	Lebanon		United Kingdom
Malawi	Mauritius		United States of America
Mali	Mexico		
Mozambique	Mongolia		
Namibia	Montenegro		
Nepal	Morocco		
Nicaragua	North Macedonia		
Niger	Oman		
Nigeria	Panama		
Pakistan	Peru		
Paraguay	Philippines		
Rwanda	Qatar		
Senegal	Republic of Moldova		
Sri Lanka	Romania		
Tajikistan	Russian Federation		
Togo	Saudi Arabia		
Uganda	Serbia		

United Republic of Tanzania	South Africa
Yemen	Thailand
Zambia	Trinidad and Tobago
Zimbabwe	Tunisia
	Turkey
	Ukraine
	Uruguay
	Viet Nam

Source: own elaboration.

The subsequent stage involves the identification of the strengths and weaknesses inherent to each cluster, based on the difference between the cluster's mean and the overall mean of each variable in the full sample. By doing so, we can unveil the indicators that contribute the most towards the differentiation between clusters. As shown in Table 12, each cluster has a specific set of indicators that are considered innovation strengths that contribute to differentiating it from other clusters. On the other hand, there are some indicators, where their means are less than the overall means representing the innovation weaknesses of the cluster.

Starting with the Specials cluster, the results reveal that this cluster's main strengths are in the field of R&D, institutions, and entrepreneurship culture. Additionally, it shows that high-quality knowledge creation is also one of the main strengths of this cluster, indicating the crucial role of synergy between the knowledge creation process and its commercialization in the business sector. As for the weaknesses, the results show that the main points are microfinance and ICT services imports. This can be attributed to the predominant focus of finance on larger-scale innovative activities within the business sector. Additionally, the limited imports of ICT services in these countries can be attributed to their high levels of ICT services productivity.

Table 12. Strengths and weaknesses of each cluster.

The Specials		The advanced		The intermediates		The primitives	
Strengths	**Weaknesses**	**Strengths**	**Weaknesses**	**Strengths**	**Weaknesses**	**Strengths**	**Weaknesses**
Gross expenditure on R&D	Loans from microfinance institutions	Graduates in science and engineering	Joint venture/strategic alliance deals/	Pupil-teacher ratio, secondary	ICT access	ICT services imports, % total trade	Venture capital received
Researchers	ICT services imports, % total trade	ICT access	PCT patents by origin	Venture capital received, value	ICT use	ICT services exports, % total trade	Pupil-teacher ratio, secondary
Firms offering formal training	Printing and other media, % manufacturing	Tertiary enrolment	State of cluster development and depth	State of cluster development and depth	School life expectancy, years	New businesses	Cost of redundancy dismissal
Citable documents H-index		Market capitalization	Gross expenditure on R&D	Cost of redundancy dismissal	PISA scales in reading, maths and science	Generic top-level domains (TLDs)	Gross capital formation
Entertainment and media market		Finance for startups and scaleups	Generic top-level domains (TLDs)	Loans from microfinance institutions	Tertiary enrolment	Patent families	Finance for startups and scaleups
Logistics performance		E-participation	GERD financed by abroad	Labour productivity growth	GERD financed by business	GDP/unit of energy use	Labor productivity growth
Global corporate R&D investors		Government's online service	Firms offering formal training	Gross capital formation	E-participation	Printing and other media	Domestic industry diversification
Rule of law		ICT use	Entertainment and media market		Government's online service	GERD financed by abroad, % GDP	Graduates in science and engineering
QS university ranking, top 3		School life expectancy, years			Domestic market scale	High-tech exports, % total trade	Utility models by origin
Entrepreneurship policies and culture		GERD financed by business, %				Mobile app creation	

Source: own elaboration.

In relation to the Advanced cluster, the findings indicate that ICT use and access, ICT-related public services, and finance for startups are the prominent strengths. This suggests a widespread integration of digital technologies within the economy and underscores the significance of digital transformation as a driving force in their innovation systems. On the other hand, PCT patent and R&D expenditure are the main weaknesses of this cluster implying the lack of knowledge commercialization in this cluster.

The leading strengths of the Intermediates cluster are mainly gross capital formation, venture capital, and labor productivity. However, this cluster also faces weaknesses in terms of lower quality in higher education, ICT infrastructure, and R&D expenditure. Consequently, it is imperative for these countries to prioritize the improvement of their ICT infrastructure, particularly in terms of affordability, as well as the enhancement of their higher education system to address these weaknesses effectively.

Regarding Primitives, their strengths are basically concentrated on ICT-related services and low-level creative activities such as media and mobile app creation. This reflects, to some extent, the low level of innovativeness in these countries, taking into consideration their limited financial resources and low-skilled labour force. On the other hand, this cluster suffers from several innovation-related issues such as the low level of finance for startups and labor productivity, in addition to the low level of education and training systems. This situation calls for creating national training and digital skills development programs to fill this gap. Moreover, policymakers should focus more on microfinance-supporting policies to overcome the resources and skills issues.

It should be borne in mind that the comparison between the means of each indicator among the four clusters can contribute to determining which indicators have a major/minimal role in the classification/ranking of the countries. Considering the aforementioned, we identify

17 indicators that have a negligible mean difference among the four clusters, see Table 13[13]. This implies the possibility of reducing the GII indicators to 63 indicators, which could capture and explain the structural differences between countries in regard to innovation performance, without altering the final conclusions. Also, this can help to reduce the number of indicators that a given country should continuously manage and monitor. This streamlined approach enhances the impact of innovation-driven activities and policymaking, enabling a more focused and effective approach to drive positive outcomes.

Table 13. Indicators that show negligible mean difference across the four clusters.

		Mean			
	Indicator	Specials	Advanced	Intermediates	Primitives
1	Cost of redundancy dismissal	13.71	14.61	19.68	21.72
2	Ease of starting a business	92.93	89.08	83.82	81.26
3	Expenditure on education	5.34	4.80	4.26	4.31
4	Government funding/pupil, secondary	23.72	23.17	20.20	18.38
5	School life expectancy	17.21	15.73	14.51	10.97
6	Pupil-teacher ratio, secondary	12.30	10.06	13.73	22.67
7	Graduates in science & engineering	24.36	23.55	24.25	18.19
8	Gross capital formation	23.60	22.35	24.89	24.72
9	GDP/unit of energy use	12.41	10.58	9.45	8.09
10	Ease of getting credit	63.89	62.50	63.14	50.38
11	Ease of protecting minority investors	67.47	63.61	63.19	49.27
12	Applied tariff rate	1.86	2.20	3.78	7.89
13	Intensity of local competition	76.43	72.92	67.58	62.82
14	Firms offering formal training	40.57	38.98	36.13	33.49
15	JV-strategic alliance deals	0.09	0.03	0.01	0.01
16	Growth rate of PPP$ GDP/worker	0.94	1.87	1.35	1.49
17	Printing & other media	1.38	2.54	1.38	1.34

Source: own elaboration.

The GII divides countries into four clusters according to one single dimension, which is the level of income (High-income, Upper-middle income, Lower-middle income, and Low income), and provides a rank of the countries included in each of these clusters. In this regard,

[13] Alternatively, in a different approach of comparison conducted in a research paper currently under review, we compared the cluster mean to the overall variable mean. This approach allowed us to pinpoint 35 indicators with a minimal mean difference across all four clusters. For the complete list of these indicators, please refer to Appendix 4.

it has to be stated that the new clustering provided above is inclusive of the 80 dimensions considered by the GII (i.e., the 80 GII indicators). Furthermore, Table 14 presents the countries ranked according to the PCA-DEA efficiency scores within their respective clusters, which every country can articulate a meaningful multi-faceted benchmark for their innovation systems. For example, in the Intermediates cluster, there are countries from different levels of income such as Qatar, Trinidad and Tobago, and Croatia from a high-income level, and Viet Nam and Jordan from a lower-middle income. This can be inferred from the efficiency of the innovation systems for these countries and shall encourage them to re-evaluate their innovation policy and strategies to improve their performance. In the case of resourceful countries such as Qatar and Trinidad and Tobago, the expansion of the innovation system is recommended for better resource exploitation. The United Arab Emirates can serve as a helpful example for them. Considering that it has gained a ranking of 35 in the efficiency PCA-DEA rank. Moreover, it has even been allocated in the Advanced cluster, despite having a similar income group and many similarities with Qatar and Trinidad and Tobago. Meanwhile, for countries with limited resources, the recommendation is to reduce the size of the innovation system and concentrate it on certain innovation activities.

Additionally, Table 14 highlights that countries such as Russia and Turkey are ranked in the second quartile based on the GII. However, when considering innovation characteristics, these countries are positioned within the Intermediates cluster. Consequently, the clustering methodology employed in this study sheds light on certain limitations of the GII in accurately capturing countries' actual innovation performance. This underscores the importance for policymakers to leverage benchmarking analysis in shaping their innovation policies.

Notably, Table 14 shows that countries such as China, Cyprus, and Estonia, fall into the second cluster (the Advanced), even though they have rankings below 25 in the PCA-DEA efficiency rank. Starting with China, this can be referred to the relatively low or moderate performance in the sub-pillars such as, but not limited to, the Regulatory environment and Business environment, including indicators like Regulatory quality, Rule of law, and Ease of resolving insolvency. In addition to, other sub-pillars such as Innovation linkages and Knowledge creation, in particular, indicators like GERD financed by abroad, JV-strategic alliance deals, Scientific and technical articles, and highly cited documents.

With respect to Cyprus and Estonia, this placement can be attributed to their relatively modest performance in various indicators including, but not limited to, Domestic market scale, Research talent % in business enterprise, highly cited documents, as well as High- & medium-high-tech manufactures. Furthermore, there are indicators in common where all of the three countries have relatively low performance, encompassing Environmental performance, Venture capital deals, Knowledge-intensive employment, and University/industry research collaboration.

Table 14. Countries ranked within their new clusters according to the PCA-DEA efficiency scores.

The Primitives			**The Intermediates**			**The Advanced**			**The Specials**		
Country	**PCA-DEA**	**GII**	**Country**	**PCA-DEA**	**GII**	**Country**	**PCA-DEA**	**GII**	**Country**	**PCA-DEA**	**GII**
Jamaica	71	81	Chile	42	51	China	9	14	Switzerland	1	1
Kenya	81	77	Thailand	43	43	Cyprus	13	28	United States of America	2	3
Namibia	89	101	Russian Federation	44	46	Estonia	23	24	Republic of Korea	3	11
Paraguay	90	95	Croatia	45	44	Italy	25	30	Germany	4	9
Rwanda	91	94	Belarus	46	72	Slovenia	26	31	Hong Kong, China	5	13
Sri Lanka	92	89	South Africa	47	63	Portugal	27	32	Netherlands	6	4
Egypt	93	92	Serbia	48	57	Spain	29	29	Japan	7	15
Honduras	94	104	Romania	49	50	Czech Republic	31	26	United Kingdom	8	5
El Salvador	95	108	Viet Nam	50	42	Malta	32	27	Sweden	10	2
Cambodia	96	98	Colombia	51	67	Bulgaria	33	40	Denmark	11	7
Botswana	97	93	Mauritius	52	82	Malaysia	34	35	Luxembourg	12	18
Guatemala	98	107	Turkey	53	49	United Arab Emirates	35	36	Finland	14	6
Nepal	99	109	Bahrain	54	78	Greece	36	41	Israel	15	10
Ecuador	100	99	Mexico	55	56	Latvia	37	34	Australia	16	22
Ghana	102	106	Uruguay	56	62	Hungary	38	33	Ireland	17	12
Pakistan	103	105	Brunei Darussalam	57	71	Slovakia	39	37	Singapore	18	8
Cote d'Ivoire	104	103	Oman	58	80	Poland	40	39	New Zealand	19	25
Kyrgyzstan	105	90	Costa Rica	59	55	Lithuania	41	38	Canada	20	17
Guinea	106	125	Qatar	60	65				Norway	21	19
Bolivia	107	110	Montenegro	61	45				France	22	16
Nigeria	108	114	India	62	52				Iceland	24	20
Tanzania	109	97	Brazil	63	66				Belgium	28	23
Senegal	110	96	Georgia	64	48				Austria	30	21
Nicaragua	111	120	Ukraine	65	47						
Uganda	112	102	Azerbaijan	66	84						

Bangladesh	113	116	Armenia	67	64
Togo	114	126	Kuwait	68	60
Zambia	115	124	Philippines	69	54
Tajikistan	116	100	Iran	70	61
Mali	117	112	Panama	72	75
Cameroon	118	115	Mongolia	73	53
Algeria	119	113	Republic of Moldova	74	58
Malawi	120	118	Tunisia	75	70
Ethiopia	121	111	Lebanon	76	88
Mozambique	122	119	Saudi Arabia	77	68
Burkina Faso	123	117	Morocco	78	74
Madagascar	124	121	Kazakhstan	79	79
Benin	125	123	Indonesia	80	85
Niger	126	127	Albania	82	83
Zimbabwe	127	122	Bosnia and Herzegovina	83	76
Burundi	128	128	Argentina	84	73
Yemen	129	129	Macedonia	85	59
			Dominican Republic	86	87
			Jordan	86	86
			Peru	87	69
			Trinidad and Tobago	88	91

Source: own elaboration.

4.5. CHAPTER SUMMARY

In this chapter, we aimed to address two key questions that arise when using CIs for measuring performance and deriving rankings: Does the ranking effectively reflect a unit's performance, and can it change with a different perspective? Additionally, how robust are these results considering the methods used? These questions emphasize the need to carefully assess CIs, taking into account their sensitivity to methodology and the potential for differing viewpoints.

Initially, PCA was applied individually to every GII sub-pillar. This process resulted in the transformation of 80 original indicators into 21 variables, with each variable representing a weighted linear combination of the indicators within the sub-pillar. Furthermore, the Benefit-of-Doubt DEA technique was applied seven times, corresponding to the seven pillars of the GII. By doing so, an overall PCA-DEA efficiency score was computed for each country, derived from the average of these seven DEAs.

Consequently, a new efficiency-based ranking was established, including all the GII 129 countries, allowing for a comparison with the original GII ranking. The analysis revealed an average absolute difference of 7.2 positions between the two rankings. Notably, countries located in the middle of the GII ranking experienced the most substantial shifts in their positions, while those situated at the extreme ends demonstrated minimal changes. To validate the robustness of the PCA-DEA results, the RF technique was employed, resulting in a congruence of 87.5% with the PCA-DEA ranking. This high level of agreement underscores the robustness of the PCA-DEA model in constructing an efficiency innovation index for GII countries.

The "Human Capital and Research" pillar has the highest average efficiency, indicating limited performance variation among countries within this pillar, as certain indicators like "School life expectancy" and "Tertiary enrollment" have limited growth potential. Conversely, the "Knowledge and Technology Outputs" pillar has the lowest average efficiency average, revealing a notable gap between inefficient countries and the efficiency frontier. This Indicates that sophisticated and Advanced domains like "Patents by origin" and "Scientific & technical articles" effectively capture performance variation among countries.

Finally, by including all the 80 GII indicators a K-means cluster analysis was applied to the GII countries, yielding a multi-dimensional innovation-driven clustering to the GII countries. As a result, the countries were grouped into four distinctive clusters. According to their descriptive statistics, these clusters were labeled as "Primitives", "Intermediates", "Advanced", and "Specials". The subsequent phase involved identifying the strengths and weaknesses inherent to each cluster by comparing the cluster mean to the overall mean of each variable across all countries. This analysis pinpointed 63 variables that significantly contributed to explaining the innovation structural differences between clusters, while 17 variables exhibited a negligible mean difference among the four clusters.

Chapter 5

CONTRIBUTIONS, LIMITATIONS AND FURTHER RESEARCH

Research is what I'm doing when I don't know what I'm doing- Wernher von Braun.

INTRODUCTION

The fundamental research questions that we aim to answer in this dissertation are twofold: Does the ranking provided by the GII effectively reflect the innovation performance of a country, or can it change when a different perspective is applied? And, how robust are the results against the methodology used in such CIs? In particular, this dissertation is centered on the CIs that are oriented to assess the innovation performance of countries, by focusing on GII. With this research, therefore, we aim to offer a novel evaluative perspective concerning countries' innovation performance. This is achieved through the introduction of an efficiency-based ranking system for the countries included in the GII. Concurrently, the purpose of the research is to complement the results offered by the GII, by categorizing the countries considered in it into distinct clusters by means of a multidimensional approach rooted in innovation-driven indicators. This not only contributes to better discerning the innovation indicators that play a pivotal role in delineating these newly-formed clusters but also provides a deeper understanding of the intricate relationships among various dimensions of innovativeness.

This chapter is structured into three sub-sections. Section 5.1 highlights the main theoretical contributions that emanate from the current research. Section 5.2 discusses the main limitations of the undertaken research. Finally, Section 5.3 introduces potential further research avenues that may help to overcome the previous limitations.

5.1. THEORETICAL CONTRIBUTION

Considering the significance of CIs in measuring various phenomena including innovation, which plays a crucial role for economic growth and development. This implies recognizing the necessity of assessing innovation performance in countries from multiple perspectives. In this regard, this dissertation contributes to the literature along several lines.

Firstly, to the best of our knowledge, the PCA-DEA model introduced in this dissertation is the first efficiency-based ranking for countries' innovation performance at global level (i.e., based on GII data). The efficiency-based ranking allows us to perceive the countries' innovation performance from an efficiency point of view, which has implications for identifying areas for improvement and policymaking. Previous research work has emphasized the multidimensional nature of innovation and the need for different perspectives in measuring innovation, such as Zabala-Iturriagagoitia et al. (2007a), Decancq and Lugo (2013), and Edquist et al. (2018).

Secondly, the level of subjectivity by human intervention in the process of generating and assigning the weights is one of the major challenges in constructing CIs (Decancq and Lugo, 2013; OECD, 2008; Greco et al., 2019). Hence, by using data-driven techniques such as PCA and DEA, the PCA-DEA model introduced in this dissertation was able to produce a robust ranking for the GII countries using intervention-free weights generation.

Thirdly, the results of this dissertation confirm the robustness of the GII ranking for both the top and bottom-ranked countries in the GII. However, it evidences that the ranking of countries with medium innovation performance is more sensitive towards the change in the weighting and aggregation schemes. Thus, what might seem significant for one group of countries may not carry the same weight for others, as highlighted in studies such as those by Zabala-Iturriagagoitia et al. (2007a) and Stavbunik & Pelucha (2019).

Fourth, despite the misleading conclusions that may be drawn due to the absence of robustness analysis for CIs, it has been often neglected (OECD, 2008). To the best of our knowledge, this dissertation is the first to introduce the RF classification as a tool to examine the robustness of CIs. Moreover, the findings of this dissertation emphasize the RF's effectiveness in forecasting a country's innovation performance. This predictive capability could potentially entail the utilization of the "important" indicators, highlighted by the RF to serve as a viable proxy for gauging and comparing countries' levels of innovation performance. This contribution is in line with what Pence et al. (2018) and Cui et al. (2020) were trying to do by performing the Artificial Neural Network.

Fifth, this research contributes to the literature by examining the structural differences and similarities in innovation among countries and categorizing them into distinct clusters according to their GII indicator performance. To the best of our knowledge, this dissertation is the first to introduce a multidimensional innovation-driven clustering for countries at global level. This expands existing work on innovation performance clustering by Cherchye et al. (2008), Roszko-Wojtowicz and Białek (2018), and Sagan (2013), Jankowska et al. (2017), Dahesh et al. (2020), Guell (2020), Lee et al. (2021), Liudmyla et al. (2022).

Sixth, the findings of this dissertation suggest that the inefficiency of a resourceful innovation system is related to the size of that system. Consequently, expanding the size of the system might result in more effective resource utilization. This aligns with the insights presented by Barbero et al. (2021), who discussed the evaluation of innovation systems by considering the quantity of available inputs and their impact on system performance. Likewise, this principle can be extended to innovation systems with limited resources, prompting them to streamline their innovation efforts toward specific areas of specialization to enhance their efficiency.

Finally, the results of this dissertation underscore the potential for reducing the dimensionality of measuring the innovation performance of countries by identifying indicators with minimal impact on capturing structural differences among countries. This contribution aligns with the findings of previous research such as Pence et al. (2018) and Cui et al. (2020).

In summary, this dissertation contributes to the existing knowledge by introducing a novel perspective for evaluating countries' innovativeness. It also examines different methodologies to determine whether the results of CIs vary based on data processing, weighting, or aggregation methods which is referred to as the robustness of the CIs. Furthermore, it pioneers a multidimensional innovation-driven clustering approach for countries on a global level. Table 15 offers a summary of the theoretical contributions of the dissertation.

Table 15. Summary of the theoretical contributions of the dissertation.

Research Questions and Goals	Results	Contributions
RQ1. Does the ranking provided by the GII effectively reflect a country’s innovation performance, and can it change with a different perspective?	A new efficiency-based ranking is established, including all the GII 129 countries, allowing for a comparison with the original GII ranking.	The dissertation contributes to the literature with a novel efficiency-based ranking for countries’ innovation performance at the global level using the PCA-DEA approach.
RQ2. How robust are the results against the methodology used in such CIs?	Countries located in the middle of the GII ranking experienced the most substantial shifts in their positions, while those situated at the extreme ends demonstrated minimal changes.	The PCA-DEA efficiency-based ranking is achieved with a congruence of 87.5% with the RF. This high level of agreement underscores the robustness of the PCA-DEA model in constructing an efficiency innovation index for GII countries.
Goal 1: · To provide a new evaluation perspective for countries’ innovativeness by introducing an efficiency-based ranking for the GII countries.	By using data-driven techniques such as PCA and DEA, we achieve a robust efficiency-based ranking for the GII countries, using intervention-free weights generation.	The results of this research evidence that the ranking of countries with medium innovation performance is more sensitive towards the change in the weighting and aggregation schemes.
Goal 2: · To group the GII countries into different clusters by using a multi-dimensional innovation-driven approach.	Countries were grouped into four distinctive clusters. According to their descriptive statistics, these clusters were labeled as "Primitives", "Intermediates", "Advanced," and "Specials".	The results indicate that the inefficiency of resourceful innovation systems is related to the size of that system. Thus, increasing the size of the system may lead to a better exploitation of the resources. Similarly, this can be applied to systems with limited resources to narrow down their innovation activities to focus on specific specializations to increase their efficiency.
Goal 3: · To reveal the innovation indicators that contribute toward the differentiation of the new clusters.	Identifying the strengths and weaknesses inherent to each cluster by comparing the cluster mean to the overall mean of each variable across all countries. This analysis pinpointed 63 variables that significantly contributed to explaining the innovation structural differences between clusters, while 17 variables exhibited a negligible mean difference among the four clusters.	This dissertation has identified GII innovation indicators that have a minimal role in measuring countries' innovation performance. It contributes to the literature by evidencing the potential of using the Cluster Analysis and the RF’s predictive capacity to serve as a viable proxy for gauging and comparing countries' levels of innovation performance.

Source: own elaboration.

5.2. LIMITATIONS

In this dissertation, certain limitations must be acknowledged.

1. This dissertation is limited to the data available for the 2019 edition of the GII. Innovation performance is dynamic, and trends can change over time. Therefore, this research would offer more informative insights if conducted over consecutive years, allowing for a dynamic assessment of national innovation performance.
2. In the context of DEA, efficiency typically refers to the ability of an entity (such as a country in the case of national innovation performance) to produce maximum output with a given set of inputs or resources. Nevertheless, in the context of innovation, where all indicators can be regarded as achievements, the Benefit-of-the-Doubt approach operates primarily within an output-centric framework. Given that this approach compares outputs with outputs, it has sparked a scholarly debate regarding whether it truly measures efficiency or effectiveness. For instance, Lafuente et al. (2022) and Lahouel et al. (2022) view the Benefit-of-the-Doubt approach as a gauge of efficiency, while Cherchye et al. (2007) consider it valuable for assessing a 'pure output setting', (equivalent to effectiveness). Meanwhile, Jovanović et al. (2022) use the terms efficiency and effectiveness interchangeably in their analysis.
3. Despite diligently implementing all the measures advised by the literature, the clustering methodology applied to categorize countries into four clusters relies on the choice of clustering algorithms and the number of clusters selected. Different choices in these aspects could lead to alternative cluster assignments and interpretations.
4. While the dissertation provides insights into innovation patterns and identifies areas of strength and weakness, the translation of these findings into effective policy

recommendations may require further in-depth analysis and consideration of contextual factors.

5.3. FURTHER RESEARCH

1. Considering the complexity and dynamic nature of innovation, future research could extend this analysis over multiple years to track changes in innovation performance and cluster dynamics over time. This would provide a more comprehensive understanding of trends and developments in national innovation systems.
2. To conduct comparative studies that apply the PCA-DEA model to different innovation indices apart from the GII. Comparing the results obtained from various indices could shed light on the robustness and consistency of the model and provide insights into the strengths and weaknesses of different CIs.
3. Future research could delve deeper into the policy implications of our findings. Investigate how specific countries have used their cluster categorization to inform and adjust their innovation policies. Are there discernible trends in policy changes based on cluster assignments? What are the implications for innovation policy design and evaluation?
4. Exploring the possibility of dynamic clustering, where countries can transition between clusters over time based on changes in their innovation performance. This could capture the evolving nature of innovation and provide more nuanced insights into how countries progress or regress in their innovation performance.
5. Consider incorporating advanced machine learning techniques for the enhancement of innovation measurement and the forecasting innovation performance. To name but a few, Deep learning analysis, Random forests analysis, and Artificial neural networks.

5.4. CHAPTER SUMMARY

This dissertation acknowledges several limitations that provide directions for further research. Firstly, it is confined to the data from the 2019 edition of the GII, limiting its temporal scope in an ever-evolving field. Future studies could offer richer insights by extending the analysis over consecutive years. Secondly, there exists a debate surrounding the Benefit-of-the-Doubt DEA approach, questioning whether it genuinely measures efficiency or effectiveness in innovation assessment. Further research could explore this debate and its implications. Thirdly, the clustering methodology used is sensitive to the choice of algorithms and cluster count, introducing subjectivity. Future work could explore alternative clustering approaches. Additionally, while the dissertation offers insights into innovation patterns, it suggests a need for more in-depth analysis to translate findings into effective policy recommendations, considering contextual factors. In light of these limitations, future research could track innovation trends over multiple years, conduct comparative analyses across different indices, delve into policy implications, explore dynamic clustering, and employ advanced machine-learning techniques for enhanced innovation measurement and prediction.

REFERNCES

Adler, N., & Golany, B. (2002). Including principal component weights to improve discrimination in data envelopment analysis. *Journal of the Operational Research Society*, *53*, 985-991.

Adler, N., & Yazhemsky, E. (2010). Improving discrimination in data envelopment analysis: PCA–DEA. *European Journal of Operational Research, 202*, 273–284.

Almirall, E., & Casadesus-Masanell, R. (2010). Open versus closed innovation: A model of discovery and divergence. *Academy of management review*, *35*(1), 27-47.

Alnafrah, I. (2021). Efficiency evaluation of BRICS's national innovation systems based on bias-corrected network data envelopment analysis. *Journal of Innovation and Entrepreneurship, 10*(1), 1-28.

Atkinson, T., Cantillon, B., Marlier, E., & Nolan, B. (2002). *Social indicators: The EU and social inclusion.* New York: Oxford University Press. doi:10.1093/0199253498.001.0001

Ausloos, M., Bartolacci, F., Castellano, N. G., & Cerqueti, R. (2018). Exploring how innovation strategies at time of crisis influence performance: a cluster analysis perspective. *Technology Analysis & Strategic Management*, *30*(4), 484-497.

Bandura, R. (2011). *Composite indicators and rankings: Inventory 2011. Technical report.* New York: United Nations Development Programme (UNDP).

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science, 30*(9), 1031-1142.

Barbero, J., Zabala-Iturriagagoitia, J. M., & Zofío, J. L. (2021). Is more always better? On the relevance of decreasing returns to scale on innovation. *Technovation, 107*, 102314.

Batool, F., & Hennig, C. (2021). Clustering with the Average Silhouette Width. *Computational Statistics & Data Analysis*, *158*, 107190. https://doi.org/10.1016/j.csda.2021.107190

Bielińska-Dusza, E., & Hamerska, M. (2021). Innovative activity of Polish enterprises–a strategic aspect. The similarity of NACE divisions. *Journal of Entrepreneurship, Management and Innovation, 17*(2), 53-98.

Billaut, J. C., Bouyssou, D., & Vincke, P. (2010). Should you believe in the Shanghai ranking? *Scientometrics, 84*(1), 237–263.

Blashfield, R. K., & Aldenderfer, M. S. (1978). The literature on cluster analysis. *Multivariate behavioral research*, *13*(3), 271-295.

Bloch, C., & Bugge, M. M. (2013). Public sector innovation—From theory to measurement. *Structural change and economic dynamics*, *27*, 133-145.

Borrás, S., & Edquist, C. (2019). *Holistic innovation policy: Theoretical foundations, policy problems, and instrument choices*. Oxford University Press.

Boschma, R. (2015) Towards an Evolutionary Perspective on Regional Resilience, Regional Studies, 49:5, 733-751, DOI: 10.1080/00343404.2014.959481

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5-32.

Breschi, S., & Malerba, F. (1997). Sectoral innovation systems: technological regimes, Schumpeterian dynamics, and spatial boundaries. *Systems of innovation: Technologies, institutions and organizations*, *1*, 130-156.

Broughel, J., & Thierer, A. (2019). *Technological innovation and economic growth: A brief report on the evidence.* Arlington: Mercatus research.

Calia, R. C., Guerrini, F. M., & Moura, G. L. (2007). Innovation networks: From technological development to business model reconfiguration. *Technovation*, *27*(8), 426-432.

Carlsson, B., & Stankiewicz, R. (1991). On the nature, function and composition of technological systems. *Journal of evolutionary economics*, *1*, 93-118.

Charles, V., Aparicio, J., & Zhu, J. (2019). The curse of dimensionality of decision-making units: A simple approach to increase the discriminatory power of data envelopment analysis. *European Journal of Operational Research*, *279*(3), 929-940.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429-444.

Cherchye, L., Moesen, W., & Van Puyenbroeck, T. (2004). Legitimately diverse, yet comparable: On synthesizing social inclusion performance in the EU. *Journal of Common Market Studies, 42*(5), 919-955.

Cherchye, L., Moesen, W., Rogge, N., & Puyenbroeck, T. V. (2007). An introduction to 'Benefit of the Doubt' composite indicators. *Social Indicators Research*, *82*(1), 111-145.

Cherchye, L., Moesen, W., Rogge, N., Van Puyenbroeck, T., Saisana, M., & Saltelli, A. (2008). Creating composite indicators with DEA and robustness analysis: The case of the Technology Achievement Index. *Journal of the Operational Research Society, 59*(2), 239–251.

Chesbrough, H. W. (2003). *Open innovation: The new imperative for creating and profiting from technology*. Harvard Business Press.

Cho, Y. J., Leem, C. S., & Shin, K. T. (2008). The relationships among manufacturing innovation, competitiveness, and business performance in the manufacturing industries of Korea. *The International Journal of Advanced Manufacturing Technology*, *38*, 840-850.

Christensen, C., & Raynor, M. (2013). *The innovator's solution: Creating and sustaining successful growth*. Harvard Business Review Press.

Cooke, P., Uranga, M. G., & Etxebarria, G. (1997). Regional innovation systems: Institutional and organisational dimensions. *Research policy*, *26*(4-5), 475-491.

Corrente, S., Bernabeu, A. G., Greco, S., & Makkonen, T. (2021). Robust measurement of innovation performances in Europe with a hierarchy of interacting composite indicators. *Economics of Innovation and New Technology*, 1-18. doi:https://doi.org/10.1080/10438599.2021.1910815

Cox, D. R., Fletcher, A. E., Gore, S. M., Spiegelhalter, D. J., & Jones, D. R. (1992). Quality of life assessment: Can we keep it simple? *Journal of the Royal Statistical Society. Series A (Statistics in Society), 155*(3), 353–393.

Crespo, N., & Crespo, C. (2016). Global innovation index: Moving beyond the absolute value of ranking with a fuzzy-set analysis. *Journal of Business Research, 69*(11), 5265-5271.

Cui, R., Sun, J., Li, Y., Yang, K., & Wu, X. (2020). Data-driven approach with artificial neural network for Global Innovation Index re-evaluation. *2020 6th International Conference on Big Data and Information Analytics (BigDIA),* (pp. 88-93).

Dahesh, M. B., Tabarsa, G., Zandieh, M., & Hamidizadeh, M. (2020). Reviewing the intellectual structure and evolution of the innovation systems approach: A social network analysis. *Technology in Society*, *63*, 101399.

Damanpour, F. (1987). The adoption of technological, administrative, and ancillary innovations: Impact of organizational factors. *Journal of management, 13*(4), 675-688.

Damanpour, F. (2014). Footnotes to research on management innovation. *Organization studies, 35(9), 1265-1285.*

Damanpour, F., & Aravind, D. (2012). Managerial innovation: Conceptions, processes and antecedents. *Management and organization review, 8(2), 423-454.*

Damanpour, F and WM Evan (1984). Organizational innovation and performance: The problem of organizational lag. Administrative Science Quarterly, 29, 392–409.

Davidson, I. (2002). Understanding K-means non-hierarchical clustering. *Computer Science Department of State University of New York (SUNY), Albany*.

Decancq, K., & Lugo, M. A. (2013). Weights in multidimensional indices of wellbeing: An overview. *Econometric Reviews, 32*(1), 7–34.

Desai , M., Fukuda-Parr, S., Johansson, C., & Sagasti, F. (2002). Measuring technology achievement of nations and the capacity to participate in the network age. *Journal of Human Development*, *1*(3), 301-311.

Despotis, D. (2005). Measuring human development via data envelopment analysis: The case of Asia and the Pacific. *Omega, 33*(5), 385-390.

Driver, H. E., & Kroeber, A. L. (1932). *Quantitative expression of cultural relationships* (Vol. 31, No. 4). Berkeley: University of California Press.

Dutta, S., Lanvin, B., & Vincent, S. W. (2019). *The Global Innovation Index 2019. Creating healthy lives—The future of medical innovation.* Ithaca, Fontainebleau, and Geneva: Cornell University, INSEAD, WIPO. From https://www.wipo.int/edocs/pubdocs/en/wipo_pub_gii_2019.pdf

Dutta, S., Lanvin, B., Wunsch-Vincent, S., & León, L. R. (Eds.). (2022). *Global Innovation Index 2022:: What is the Future of Innovation-driven Growth?* (Vol. 2000). WIPO.

Dziallas, M., & Blind, K. (2019). Innovation indicators throughout the innovation process: An extensive literature analysis. *Technovation*, *80*, 3-29.

Edquist, C. (Ed.). (1997). *Systems of innovation: technologies, institutions, and organizations*. Psychology Press.

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

Edquist, C., Iturriagagoitia, J. Z., & Barbero, J. (2018). On the meaning of innovation performance: Is the synthetic indicator of the Innovation Union Scoreboard flawed? *Research Evaluation, 27*(3), 196–211.

Esty, D. C., Levy, M., Srebotnjak, T., & De Sherbinin, A. (2005). *Environmental sustainability index: Benchmarking national environmental stewardship.* New Haven: Yale Center for Environmental Law & Policy.

European Commission. (2019, June 17). *European Commission.* Retrieved March 30, 2020 from https://ec.europa.eu/commission/presscorner/detail/en/ip_19_2991

European Commission (2023). European Innovation Scoreboard 2023. European Commission. Directorate-General for Research and Innovation. Brussels.

ForumIAS. (2023, September 13). Retrieved from Forumias: https://blog.forumias.com/global-innovation-index-2021/

Frades, I., & Matthiesen, R. (2010). Overview on techniques in cluster analysis. *Bioinformatics methods in clinical research*, 81-107.

Freudenberg, M. (2003). Composite indicators of country performance: A critical assessment. *OECD Science, Technology and Industry Working Papers*.

Gadrey, J., & Jany, F.-C. (2003). *Les indicateurs de richesee et de developpement: Un bilan international en vue d'une initiative francaise.* Paris: DARES, Ministere du Travail.

Gault, F. (2018). Defining and measuring innovation in all sectors of the economy. *Research policy*, *47*(3), 617-622.

Godin, B. (2006). The linear model of innovation: The historical construction of an analytical framework. *Science, Technology, & Human Values*, *31*(6), 639-667.

Greco, S., Ishizaka, A., Tasiou, M., & Torrisi, G. (2019). On the methodological framework of composite. *Social Indicators Research, 141*, 61–94.

Grupp, H., & Mogee, M. E. (2004). Indicators for national science and technology policy: How robust are composite indicators? *Research Policy, 33*(9), 1373-1384.

Grupp, H., & Schubert, T. (2010). Review and new evidence on composite innovation indicators for evaluating national performance. *Research Policy*, *39*(1), 67-78.

Guaita Martínez, J., Martín Martín, ,., Ostos, M. S., & Pardo, M. d. (2021). Constructing knowledge economy composite indicators using an MCA-DEA approach. *Economic Research-Ekonomska Istraživanja, 34*(1), 331-351.

Guell, F. (2020, June 20). *European innovation: EIS 2020*. Retrieved December 21, 2022, from Agile Innovation: https://www.fguell.com/en/european-innovation-eis-2020

Halliday, S. V. (2008). The power of myth in impeding service innovation: A perspective gained from analysis of service providers' narratives. *Journal of Management Inquiry*, *17*(1), 44-55.

Hamidi, S., & Berrado, A. (2018). Segmentation of innovation determinants: Case of the Global Innovation Index. *12th International Conference on Intelligent Systems: Theories and Applications* (pp. 1-8). Rabat: Association for Computing Machinery. From https://doi.org/10.1145/3289402.3289548

Harding, R. (2003). New challenges for innovation systems: a cross-country comparison. *International Journal of Technology Management*, *26*(2-4), 226-246.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning* (2 ed., Vol. 12). New York: Springer-Verlag. From https://hastie.su.domains/ElemStatLearn/printings/ESLII_print12_toc.pdf

Hatefi, M. S., & Torabi, S. A. (2010). A common weight MCDA–DEA approach to construct composite indicators. *Ecological Economics, 70*(1), 114-120.

Henderson, D. W. (1974). *Social indicators: A rationale and research framework.* Ottawa: Cambridge University Press. From https://www.cambridge.org/core/journals/canadian-journal-of-political-science-revue-canadienne-de-science-politique/article/abs/dw-henderson-social-indicators-a-rationale-and-research-framework-ottawa-information-canada-1974-pp-vii-90/8DE50AF54F6D3452E0C7

Hermans, E., Bossche, F. V., & Wets, G. (2008). Combining road safety information in a performance index. *Accident Analysis and Prevention, 40*, 1337–1344.

Herrmann, A. M., & Peine, A. (2011). When 'national innovation system'meet 'varieties of capitalism'arguments on labour qualifications: On the skill types and scientific knowledge needed for radical and incremental product innovations. *Research Policy, 40(5), 687-701.*

Ho, L. A. (2011). Meditation, learning, organizational innovation and performance. *Industrial Management & Data Systems*, *111*(1), 113-131.

Hollenstein, H. (2003). Innovation modes in the Swiss service sector: a cluster analysis based on firm-level data. *Research policy*, 32(5), 845-863.

Huizingh, E. K. (2011). Open innovation: State of the art and future perspectives. *Technovation*, *31*(1), 2-9.

Jankowska, B., Jedrych, A., & Dabrowska, K. (2017). Efficiency of national innovation systems – Poland and Bulgaria in the context of the Global Innovation Index. *Comparative Economic Research, 20*(3), 77-94.

Jovanović, M., Savić, G., Cai, Y., & Levi-Jakšić, M. (2022). Towards a Triple Helix based efficiency index of innovation systems. *Scientometrics*, *127*(5), 2577-2609.

Kawamoto, T., & Kabashima, Y. (2017). Cross-validation estimate of the number of clusters in a network. *Scientific Reports*, *7*(1), Article 1. https://doi.org/10.1038/s41598-017-03623-x

Kimberly, J. R., & Evanisko, M. J. (1981). Organizational innovation: The influence of individual, organizational, and contextual factors on hospital adoption of technological and administrative innovations. *Academy of management journal*, *24*(4), 689-713.

Kline, S. J. (1985). Innovation is not a linear process. *Research management*, *28*(4), 36-45.

Kline, S. J., & Rosenberg, N. (2010). An overview of innovation. *Studies on science and the innovation process: Selected works of Nathan Rosenberg*, 173-203.

Krishna, T. S., Yesu Babu, A., & Kiran Kumar, R. (2018). Determination of optimal clusters for a Non-hierarchical clustering paradigm K-Means algorithm. In *Proceedings of International Conference on Computational Intelligence and Data Engineering: ICCIDE 2017* (pp. 301-316). Springer Singapore.

Lafuente, E., Araya, M., & Leiva, J. C. (2022). Assessment of local competitiveness: A composite indicator analysis of Costa Rican counties using the 'Benefit of the Doubt'model. *Socio-Economic Planning Sciences, 81,* 100864.

Lahouel, B. B., Zaied, Y. B., Taleb, L., & Kočišová, K. (2022). The assessment of socio-environmental performance change: A Benefit of the Doubt indicator based on Directional Distance Function and Malmquist productivity index. *Finance Research Letters*, *49*, 103164.

Lavopa, A., & Szirmai, A. (2018). Structural modernisation and development traps. An empirical approach. *World Development*, *112,* 59-73.

Lee, K., Lee, J., & Lee, J. (2021). Variety of national innovation systems (NIS) and alternative pathways to growth beyond the middle-income stage: Balanced, imbalanced, catching-up, and trapped NIS. *World Development*, *144*, 105472.

Limoges, C., Scott, P., Schwartzman, S., Nowotny, H., & Gibbons, M. (1994). The new production of knowledge: The dynamics of science and research in contemporary societies. *The New Production of Knowledge*, 1-192.

Liudmyla, V., Svitlana, T., Olha, M., Mariia, R., & Viktor, T. (2022). Comparative Analysis of Strategies for Innovative Development of The Economy: The Experience of The EU Countries. *Quality-Access to Success, 23(189).*

Lundvall, B. (Editor), 1992, National Systems of Innovation. Towards a Theory of Innovation and Interactive Learning.

Lundvall, B. A. (2007). National innovation systems—Analytical concept and development tool. *Industry and Innovation, 14*(1), 95-119.

Ma, X., Gryshova, I., Khaustova, V., Reshetnyak, O., Shcherbata, M., Bobrovnyk, D., & Khaustov, M. (2022). Assessment of the Impact of Scientific and Technical Activities on the Economic Growth of World Countries. *Sustainability, 14*(21), 14350.

Martin, J., Mendoza, C., & Román , C. (2017). A DEA Travel–Tourism Competitiveness Index . *Social Indicators Research, 130*, 937–957.

Melyn, W., & Moesen, W. (1991). *Towards a synthetic indicator of macroeconomic performance.* Leuven: Center for Economic Studies.

Metcalfe, J. S. (1995). Technology systems and technology policy in an evolutionary framework. *Cambridge journal of economics*, *19*(1), 25-46.

Moldan, B., & Billharz, S. (1997). *Sustainability indicators : A report on the project on indicators of sustainable development.* New York: Chichester John Wiley & Sons.

Munda, G. (2005). Multiple criteria decision analysis and sustainable development. In S. Greco , M. Ehrgott , & J. Figueira , *Multiple criteria decision analysis. International series in operations research & management science* (Vol. 233, pp. 1235-1267). New York: Springer. From https://doi.org/10.1007/978-1-4939-3094-4_27

Munda, G. (2012). Choosing aggregation rules for composite indicators. *Social Indicators Research, 109*(3), 337–354.

Munda, G., & Nardo, M. (2005). *Non-compensatory composite indicators for ranking countries: A defensible setting-—EUR 21833 EN.* Ispra: Institute for the Protection and Security of the Citizen.

Murtagh, F., & Legendre, P. (2014). Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *Journal of Classification*, *31*(3), 274–295. https://doi.org/10.1007/s00357-014-9161-z.

Naghizadeh, A., & Metaxas, D. N. (2020). Condensed Silhouette: An Optimized Filtering Process for Cluster Selection in K-Means. *Procedia Computer Science, 176*, 205–214. https://doi.org/10.1016/j.procs.2020.08.022.

Nikolaidis, Y., Fouskas, K., & Carayannis , E. G. (2013). Assisting regional policy by rapidly comparing enterprise innovation between regions. *Journal of Innovation and Entrepreneurship, 2*(1), 1-25.

Nowotny, H., Scott, P., & Gibbons, M. (2003). Introduction:'Mode 2'revisited: The new production of knowledge. *Minerva*, *41*(3), 179-194.

OECD. (2008). *Handbook on constructing composite indicators: Methodology and user guide.* Paris: OECD.

OECD. (2010). The OECD innovation strategy: innovation to strengthen growth and address global and social challenges.

OECD/Eurostat. (2018). *Oslo Manual: Guidelines for Collecting, Reporting and Using Data on Innovation, Fourth Edition.* Paris: OECD Publishing.

Oke, A., Burke, G., & Myers, A. (2007). Innovation types and performance in growing UK SMEs. *International Journal of Operations & Production Management*, *27*(7), 735-753.

Omer, M., El-Amin, M., Nasr, A., & Ahmed, R. (2020). Modeling, visualization, and analysis of African innovation performance. *International Conference on Learning Representation (ICLR).* Addis Ababa. Retrieved 3 11, 2022 from https://www.researchgate.net/publication/343735079_Modeling_Visualization_and_Analysis_of_African_Innovation_Performance/fulltext/5f3c9c16a6fdcccc43d31df6/Modeling-Visualization-and-Analysis-of-African-Innovation-Performance.pdf?_sg%5B0%5D=CEDIy0xN-y8EZs6

Omrani, H., Valipour, M., & Mamakani, S. J. (2019). Construct a composite indicator based on integrating common weight data envelopment analysis and principal component analysis models: An application for finding development degree of provinces in Iran. *Socio-Economic Planning Sciences, 68*, 100618.

Osberg, L., & Sharpe, A. (2002). An index of economic well–being for selected OECD countries. *The Review of Income and Wealth*, *48*(3), 291-316.

Panek, T. (2015). Hierarchiczny model pomiaru jakości życia. *Wiadomości Statystyczne. The Polish Statistician*, *60*(06), 1-22.

Paruolo, P., Saisana, M., & Saltelli, A. (2013). Ratings and rankings: voodoo or science? *Journal of the Royal Statistical Society: Series A (Statistics in Society) , 176*(3), 609–634.

Pence, İ., Kalkan, A., & Cesmeli, M. (2019). Estimation of the country ranking scores on the Global Innovation Index 2016 using the artificial neural network method. *International Journal of Innovation and Technology Management, 16*(4), 1940007.

Pennoni, F., Tarantola, S., & Latvala, A. (2005). *The European e-business readiness index.* Joint Research Centre (JRC). Retrieved March 20, 2020 from https://www.researchgate.net/publication/46446317_The_2005_European_e-Business_Readiness_Index

Phillips, W., Lee, H., Ghobadian, A., O'regan, N., & James, P. (2015). Social innovation and social entrepreneurship: A systematic review. *Group & Organization Management*, *40*(3), 428-461.

Qian, C., Cao, Q., & Takeuchi, R. (2013). Top management team functional diversity and organizational innovation in China: The moderating effects of environment. *Strategic Management Journal*, *34*(1), 110-120.

Ramanathan, R. (2006). Evaluating the comparative performance of countries of the Middle East and North Africa: A DEA application. *Socio-Economic Planning Sciences, 40*(2), 156-167.

Roszko-Wójtowicz, E., & Białek, J. (2018). Diverse approaches to the multidimensional assessment of innovation in the European Union. *Acta Oeconomica*, *68*(4), 521-547.

Rothwell, R. (1992). Developments towards the fifth-generation model of innovation. *Technology Analysis & Strategic Management, 4(1), 73-75.*

Sagan, A. (2013). Paradygmaty marketingu a modelowanie zmiennych ukrytych–porównanie modeli pomiarowych skali WSAW. *Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie*, *909*(09), 17-27.

Saisana, M., Saltelli, A., Schulze, N., Tarantola, S., & Duchene, V. (2005). Uncertainty and sensitivity analysis for the Knowledge-based Economy Index. *Conference on Medium-Term Economic Assessment (CMTEA).* Sofia. Retrieved 3 11, 2022 from https://www.researchgate.net/profile/Michaela-Saisana/publication/242180137_Uncertainty_and_Sensitivity_Analysis_for_the_Knowledge-Based_Economy_Index/links/579723b708aed51475e66de4/Uncertainty-and-Sensitivity-Analysis-for-the-Knowledge-Based-Economy-Inde

Saisana , M., & Tarantola, S. (2002). *State of the art report on current methodologies and practices for composite indicator development.* Ispra, Italy: European Commission, Joint Research Centre, Institute for the Protection and the Security of the Citizen, Technological and Economic Risk Management Unit.

Saltelli, A. (2007). Composite indicators between analysis and advocacy. *Social Indicators Research, 81*(1), 65–77.

Samara, E., Georgiadis, P., & Bakouros, I. (2012). The impact of innovation policies on the performance of national innovation systems: A system dynamics analysis. *Technovation*, *32*(11), 624-638.

Setiawan, B., Djanali, S., Ahmad, T., & Aziz, N. M. (2019). Assessing centroid-based classification models for intrusion detection system using composite indicators. *Procedia Computer Science*, *161*, 665-676.

Shen, Y., Hermans, E., Brijs, T., & Wets, G. (2013). Data envelopment analysis for composite indicators:. *Social Indicators Research, 114*, 739–756.

Sorzano, C. O. S., Vargas, J., & Montano, A. P. (2014). A survey of dimensionality reduction techniques. *arXiv preprint arXiv:1403.2877*.

Solow, R. M. (1957). Technical change and the aggregate production function. *The review of Economics and Statistics*, *39*(3), 312-320.

Spieth, P., Schneckenberg, D., & Ricart, J. E. (2014). Business model innovation–state of the art and future challenges for the field. *R&d Management*, *44*(3), 237-247.

Stavbunik, Y., & Pelucha, M. (2019). The innovation system of Kazakhstan in the context of the Global Innovation Index. *The Innovation Journal: The Public Sector Innovation Journal, 24*(1), 1-24.

Stefko, R., Gavurova, B., Rigelsky, M., & Ivankova, V. (2019). Evaluation of selected indicators of patient satisfaction and economic indices in OECD country. *Economics & Sociology*, *12*(4), 149-332.

Ueda, T., & Hoshiai, Y. (1997). Application of principal component analysis for parsimonious summarization of DEA inputs and/or outputs. *Journal of the Operations Research Society of Japan*, *40*(4), 466-478.

Thanassoulis, E. (2001). *Introduction to the theory and application of data envelopment analysis* (1 ed.). Dordrecht: Kluwer Academic Publishers. doi:10.1007/978-1-4615-1407-7

UNDP. (2019). *Human development index.* New York: United Nations Development Programme. Retrieved 10 19, 2020 from https://hdr.undp.org/en/content/download-data

Van de Ven, A. H., & Rogers, E. M. (1988). Innovations and organizations: Critical perspectives. *Communication research*, *15*(5), 632-651.

Vyas, V. (2005). Imitation, incremental innovation and climb down: A strategy for survival and growth of new ventures. *The Journal of Entrepreneurship*, *14*(2), 103-116.

Witell, L., Snyder, H., Gustafsson, A., Fombelle, P., & Kristensson, P. (2016). Defining service innovation: A review and synthesis. *Journal of Business Research*, *69*(8), 2863-2872.

Yang, F.-C., Kao, R.-H., Chen, Y.-T., Ho, Y.-F., Cho, C.-C., & Huang, S.-W. (2018). A common weight approach to construct composite indicators: The evaluation of fourteen emerging markets. *Social Indicators Research, 137*(2), 463-479.

Yang, J., & Lan, H. (2010, August). The moderating role of organizational learning in the relationship between organizational innovation and performance. In *2010 International Conference on Management and Service Science* (pp. 1-4). IEEE.

Yoon, D. (2017). The regional-innovation cluster policy for R&D efficiency and the creative economy: With focus on Daedeok Innopolis. *Journal of Science and Technology Policy Management*, *8*(2), 206-226.

Zabala-Iturriagagoitia, J. M., Jiménez-Sáez, F., Castro-Martínez , E., & Gutiérrez-Gracia , A. (2007a). What indicators do (or do not) tell us about regional innovation systems. *Scientometrics*, *41*(5), 661-672.

Zabala-Iturriagagoitia, J. M., Voigt, P., Gutierrez-Gracia, A., & Jimenez-Saez, F. (2007b). Regional innovation systems: How to assess performance. *Regional Studies, 41*(5), 661-672.

Zofio, J. L., Aparicio, J., Barbero, J., & Zabala-Iturriagagoitia, J. M. (2023). The influence of bottlenecks on innovation systems performance: Put the slowest climber first. *Technological Forecasting and Social Change*, *193*, 122607.

APPENDICES

Appendix 1: The GII profile for Spain that include all the pillar, sub-pillar, and the indicators.

SPAIN

GII 2019 rank

29

Output rank	Input rank	Income	Region	Population (mn)	GDP, PPP$	GDP per capita, PPP$	GII 2018 rank
28	25	High	EUR	46.4	1,867.9	40,138.8	28

		Score/Value	Rank	
	INSTITUTIONS	**78.1**	**30**	
1.1	**Political environment**	**73.5**	**33**	
1.1.1	Political and operational stability*	77.2	44	
1.1.2	Government effectiveness*	71.6	29	
1.2	**Regulatory environment**	**77.9**	**34**	
1.2.1	Regulatory quality*	67.2	34	
1.2.2	Rule of law*	73.1	30	
1.2.3	Cost of redundancy dismissal, salary weeks	17.4	74	○
1.3	**Business environment**	**83.0**	**25**	
1.3.1	Ease of starting a business*	86.9	69	○
1.3.2	Ease of resolving insolvency*	79.1	18	
	HUMAN CAPITAL & RESEARCH	**47.0**	**26**	
2.1	**Education**	**54.3**	**46**	
2.1.1	Expenditure on education, % GDP	4.3	71	○
2.1.2	Government funding/pupil, secondary, % GDP/cap	18.5	57	○
2.1.3	School life expectancy, years	17.9	13	●
2.1.4	PISA scales in reading, maths, & science	491.4	27	
2.1.5	Pupil-teacher ratio, secondary	11.6	46	
2.2	**Tertiary education**	**41.2**	**33**	
2.2.1	Tertiary enrolment, % gross	91.2	6	● ◆
2.2.2	Graduates in science & engineering, %	23.9	34	
2.2.3	Tertiary inbound mobility, %	2.7	68	○
2.3	**Research & development (R&D)**	**45.5**	**21**	
2.3.1	Researchers, FTE/mn pop.	2,873.4	32	
2.3.2	Gross expenditure on R&D, % GDP	1.2	31	
2.3.3	Global R&D companies, avg. exp. top 3, mn US$	74.2	14	●
2.3.4	QS university ranking, average score top 3*	47.0	23	
	INFRASTRUCTURE	**63.1**	**10**	●
3.1	**Information & communication technologies(ICTs)**	**87.4**	**17**	
3.1.1	ICT access*	80.5	25	
3.1.2	ICT use*	77.2	23	
3.1.3	Government's online service*	93.8	16	
3.1.4	E-participation*	98.3	5	● ◆
3.2	**General infrastructure**	**43.2**	**36**	
3.2.1	Electricity output, kWh/mn pop.	5,853.6	35	
3.2.2	Logistics performance*	82.7	17	
3.2.3	Gross capital formation, % GDP	21.8	77	○
3.3	**Ecological sustainability**	**58.8**	**8**	●
3.3.1	GDP/unit of energy use	12.5	26	
3.3.2	Environmental performance*	78.4	12	●
3.3.3	ISO 14001 environmental certificates/bn PPP$ GDP	7.3	13	● ◆
	MARKET SOPHISTICATION	**59.5**	**18**	
4.1	**Credit**	**55.3**	**24**	
4.1.1	Ease of getting credit*	60.0	66	○
4.1.2	Domestic credit to private sector, % GDP	105.7	22	
4.1.3	Microfinance gross loans, % GDP	n/a	n/a	
4.2	**Investment**	**44.6**	**58**	
4.2.1	Ease of protecting minority investors*	70.0	27	
4.2.2	Market capitalization, % GDP	63.5	26	
4.2.3	Venture capital deals/bn PPP$ GDP	0.0	29	
4.3	**Trade, competition, & market scale**	**78.6**	**14**	● ◆
4.3.1	Applied tariff rate, weighted avg., %	1.8	23	
4.3.2	Intensity of local competition*	75.8	22	
4.3.3	Domestic market scale, bn PPP$	1,867.9	15	● ◆

		Score/Value	Rank	
	BUSINESS SOPHISTICATION	**38.7**	**37**	
5.1	**Knowledge workers**	**52.1**	**34**	
5.1.1	Knowledge-intensive employment, %	33.2	40	
5.1.2	Firms offering formal training, % firms	n/a	n/a	
5.1.3	GERD performed by business, % GDP	0.7	32	
5.1.4	GERD financed by business, %	46.7	33	
5.1.5	Females employed w/advanced degrees, %	22.1	19	
5.2	**Innovation linkages**	**26.5**	**60**	
5.2.1	University/industry research collaboration*	42.2	59	
5.2.2	State of cluster development*	54.4	36	
5.2.3	GERD financed by abroad, %	8.1	47	
5.2.4	JV-strategic alliance deals/bn PPP$ GDP	0.0	55	
5.2.5	Patent families 2+ offices/bn PPP$ GDP	0.6	32	
5.3	**Knowledge absorption**	**37.5**	**46**	
5.3.1	Intellectual property payments, % total trade	1.2	28	
5.3.2	High-tech imports, % total trade	6.8	74	○
5.3.3	ICT services imports, % total trade	1.6	38	
5.3.4	FDI net inflows, % GDP	1.9	81	○
5.3.5	Research talent, % in business enterprise	37.2	35	
	KNOWLEDGE & TECHNOLOGY OUTPUTS	**37.2**	**24**	
6.1	**Knowledge creation**	**34.2**	**25**	
6.1.1	Patents by origin/bn PPP$ GDP	2.2	41	
6.1.2	PCT patents by origin/bn PPP$ GDP	0.8	31	
6.1.3	Utility models by origin/bn PPP$ GDP	1.3	20	
6.1.4	Scientific & technical articles/bn PPP$ GDP	20.0	25	
6.1.5	Citable documents H-index	59.3	12	●
6.2	**Knowledge impact**	**48.5**	**18**	
6.2.1	Growth rate of PPP$ GDP/worker, %	0.5	74	○
6.2.2	New businesses/th pop. 15-64	3.2	39	
6.2.3	Computer software spending, % GDP	0.7	6	● ◆
6.2.4	ISO 9001 quality certificates/bn PPP$ GDP	18.0	18	
6.2.5	High- & medium-high-tech manufactures, %	0.4	28	
6.3	**Knowledge diffusion**	**28.9**	**32**	
6.3.1	Intellectual property receipts, % total trade	0.5	27	
6.3.2	High-tech net exports, % total trade	3.9	36	
6.3.3	ICT services exports, % total trade	2.9	35	
6.3.4	FDI net outflows, % GDP	3.7	17	
	CREATIVE OUTPUTS	**39.7**	**31**	
7.1	**Intangible assets**	**56.7**	**14**	●
7.1.1	Trademarks by origin/bn PPP$ GDP	52.8	46	
7.1.2	Industrial designs by origin/bn PPP$ GDP	14.2	7	● ◆
7.1.3	ICTs & business model creation*	74.3	22	
7.1.4	ICTs & organizational model creation*	63.4	34	
7.2	**Creative goods & services**	**21.2**	**54**	
7.2.1	Cultural & creative services exports, % total trade	1.0	28	
7.2.2	National feature films/mn pop. 15-69	7.3	25	
7.2.3	Entertainment & Media market/th pop. 15-69	28.2	24	
7.2.4	Printing & other media, % manufacturing	1.3	41	
7.2.5	Creative goods exports, % total trade	0.9	46	
7.3	**Online creativity**	**24.0**	**30**	
7.3.1	Generic top-level domains (TLDs)/th pop. 15-69	28.0	22	
7.3.2	Country-code TLDs/th pop. 15-69	16.4	30	
7.3.3	Wikipedia edits/mn pop. 15-69	58.8	17	
7.3.4	Mobile app creation/bn PPP$ GDP	12.0	38	

NOTES: ● indicates a strength; ○ a weakness; ◆ an income group strength; ◇ an income group weakness; * an index; † a survey question. ⓘ indicates that the economy's data are older than the base year; see Appendix II for details, including the year of the data, at http://globalinnovationindex.org. Square brackets [ ] indicate that the data minimum coverage (DMC) requirements were not met at the sub-pillar or pillar level.

Source: The GII 2019 report.

Appendix 2: The 21 variables yielded from the linear combination V_q, and eventually used as output variables for the sub-DEA.

Country	Pillar 1			Pillar 2			Pillar 3			Pillar 4			Pillar 5			Pillar 6			Pillar 7		
	V1	**V2**	**V3**	**V4**	**V5**	**V6**	**V7**	**V8**	**V9**	**V10**	**V11**	**V12**	**V13**	**V14**	**V15**	**V16**	**V17**	**V18**	**V19**	**V20**	**V21**
Albania	119.31	90.32	133.88	394.6	60.57	143.02	234.84	2,320.42	68.39	77.9	112.58	76.31	28.93	66.37	20.53	3.48	6.91	1.47	98.1	11.4	26.68
Algeria	81.43	40.23	107.19	355.75	61.3	751.5	129.49	1,505.58	57.34	26.58	84.42	415.99	29.08	60.19	4.17	10.46	1.84	0.21	85.82	1.24	4.74
Argentina	117.86	85.58	103.73	450.73	84.67	1,208.13	250.37	2,876.45	59.53	52.33	81.54	557.54	45.47	69.42	13.15	26.45	6.26	2.72	124.93	12.38	21.75
Armenia	105.07	94.46	118.05	394.79	54.66	806.62	220.89	2,143.91	60.07	90.48	101.26	78.06	66.33	73.88	19.31	25.73	2.92	3.46	143.26	12.02	98.86
Australia	168.93	183.15	147.69	481.37	112.24	4,289.83	327.47	8,927.17	73.28	175.11	135.18	792.58	122.46	96.86	24.86	70.19	14.53	1.63	156.02	45.38	142.82
Austria	165.44	171.77	135.31	470.33	98.73	5,068.79	302.75	6,589.73	79.14	104.74	96.84	321.17	123.86	123.2	41.92	58	7.86	4.43	152.92	46.75	134.03
Azerbaijan	108.56	71.04	134.64	396.17	46.42	756.65	242.21	2,202.56	62.39	75.17	113.29	151.79	59.28	98.5	20.94	6.09	2.27	5.28	123.44	7.41	25.77
Bahrain	118.44	113.44	113	446.37	53.75	341.6	289.93	16,964.21	52.94	89.47	114.43	103.81	50.98	95.75	3.73	3.42	7.24	0.93	113.33	12.53	19.45
Bangladesh	81.52	61.79	91.78	369.24	22.44	314.05	196.34	359.6	34.93	55.57	90.52	480.8	29.77	66.89	14.89	10.33	2.01	0.84	90.14	1.32	2.64
Belarus	105.94	55.44	122.93	439.43	95.87	1,920.76	286.21	3,020.04	62.85	60.26	109.35	161.53	105.89	83.48	24.59	14.05	17.22	3.69	117.71	11.75	73.15
Belgium	152.07	160.47	148.95	481.56	78.89	4,600.66	283.65	6,446.00	75.37	98.29	126.88	368.31	134.12	123.67	37.04	62.89	7.42	5.73	154.99	42.21	90.88
Benin	93.83	62.34	110.55	347.55	29.04	69.59	114.72	58.47	36.98	39.82	82.9	79.05	11.75	59.3	6.52	7.8	2.49	0.34	82.98	1.82	2.46
Bolivia	84.53	41.81	89.76	399.86	53.2	152.08	188.15	747.82	55.28	82.32	87.73	105.33	27.9	52.18	5.58	6.28	3.57	0.8	86.55	4.4	6.01
Bosnia and Herzegovina	96.06	80.02	107.27	401.48	59.12	430.82	184	4,316.94	44.42	89.28	99.02	77.98	51.9	59.28	8.2	9.74	19.51	2.62	83.43	8.46	41.71
Botswana	137.63	119.77	104.58	410.6	35.12	164.71	143.88	1,044.99	54.62	64.3	124.99	73.78	40.42	64.17	4.36	7	9.1	1.83	87.39	5.63	6.75
Brazil	104	82.13	108.37	387.16	53.67	902.6	286.76	2,405.49	60.8	82.27	99.47	1,936.46	71.55	83.23	25.45	36.25	5.13	3.01	126.18	6.7	22.46
Brunei Darussalam	163.11	123.56	126.3	455.62	50.32	2,905.22	257.2	8,660.61	63.36	103.14	135.13	69.11	96.72	71.39	29.94	3.7	4.66	3.16	96.57	24.19	30.2
Bulgaria	118.54	104.1	120.32	421.94	73.75	1,954.30	275.45	5,358.02	72.72	86.25	87.35	143.77	87.68	81.34	33.2	19.21	30.44	4.45	152.84	16.6	67.83
Burkina Faso	86.66	68.14	108.7	332.76	17.27	43.74	150.21	256.32	42.36	46.34	108.59	74.61	15.59	58.59	11.62	7.11	2.44	1.11	82.85	1.1	1.31
Burundi	53.81	35.5	105.59	317.78	18.02	78.09	84.29	261.44	28.24	19.25	151.34	47.99	4.77	66.03	8.54	2.78	1.76	0.42	66.56	2.46	1.32
Cambodia	102.16	54.37	85.21	378.46	22.61	27.91	108.47	324.44	43.24	126.77	95.14	94.04	21.94	83.35	12.6	5.29	3.63	1.08	119.14	3.99	2.57
Cameroon	81.23	47.19	103.48	361.96	31.42	186.24	116.62	327.74	41.61	54.78	75.24	112.72	19.47	67.5	9.32	9.51	2.05	1.4	87.19	2.8	4.09
Canada	179.78	184.1	151.31	495.43	74.71	4,048.41	312.84	15,664.34	67.93	140.67	161.18	1,083.92	98.29	114.94	42.71	78.71	4.26	7.36	165.18	40.71	143.06
Chile	144.32	158.13	125.46	429.65	88.88	494.6	279.91	3,671.64	59.56	126.86	127.48	326.14	63.21	79.13	27.13	26.57	11.56	3.67	150.67	11.59	28.38
China	129.1	87.47	125.71	486.09	59.24	1,289.51	274.22	3,876.73	54.05	163.53	115.6	14,035.85	131.14	104.15	48.26	119.3	14.97	12.51	241	6.67	13.97
Colombia	104	92.59	128.57	404.21	66.01	110.42	262.96	1,379.67	70.69	107.03	103.66	477.05	70.09	78.05	10.63	16.05	13.48	2.14	119.01	5.35	21.48

Costa Rica	119.9	116.94	96.32	404.19	55.88	500.39	252.61	1,932.00	70.43	109.45	67.97	109.23	36.87	84.92	18.88	11.14	4.97	6.94	163.63	10.47	21.27
Côte d'Ivoire	89.95	65.76	119.31	390.63	21.4	348.47	100.56	410.65	44.43	71.43	120.65	123.12	22.74	50.67	14.36	5.92	3.59	1.33	107.12	2.26	5.04
Croatia	136.08	111.48	116.87	449.44	72.99	1,711.83	261.3	2,612.32	71.07	84.19	100.93	106.91	89.01	53.16	18.76	27.12	19.48	3.42	123.7	24.42	52.47
Cyprus	145.87	137.79	140.36	427.38	67.68	1,075.14	293.96	4,926.65	78.01	196.92	84.96	81.43	84.98	79.66	54.64	25.55	28.55	47.96	143.9	26.05	181.94
Czech Republic	150.71	154.67	137.84	469.73	73.47	3,400.33	247.24	6,951.11	81.01	90.6	113.68	283.37	94.6	93.9	42.01	39.75	23.98	9.81	151.77	19.97	97.26
Denmark	178.42	177.93	149.54	487.76	85.39	7,370.44	342.54	4,556.31	84.99	178.23	126.87	224.41	137.95	120.01	42.86	74.49	12.17	7.8	162.33	57.89	190.05
Dominican Republic	102.45	84.71	101.8	336.45	57.49	529.55	205.24	1,574.37	67.47	54.79	96.68	163.82	45.38	70.71	14.99	1.88	2.65	1.44	119.72	6.8	8.92
Ecuador	89.08	56.39	80.83	399.78	48.52	379	215.02	1,450.49	58.98	57.77	93.67	171.71	24.1	63.85	12.93	10	6.58	0.57	116.3	6.04	8.26
Egypt	85.88	66.56	106.43	382.49	36.67	632.36	181.77	1,756.33	61.81	69.16	81.06	774.39	37.04	74.65	9.15	18.21	3.57	1	110.67	1.73	4.25
El Salvador	97.03	70.19	104.41	402.12	39.5	60.39	192.26	822.37	55.34	97.99	79.22	90.04	47.64	51.61	13.2	1.51	4.25	2.68	130.76	3.15	7.65
Estonia	157.3	166	132.87	493.74	80.53	3,285.80	316.11	8,510.73	70.64	102	121.19	90.16	110.24	84.5	27.45	31.84	27.34	6.13	181.23	42.67	201.99
Ethiopia	82.4	56.55	85.29	311.41	22.03	41.74	141.81	112.75	41.65	26.09	267.15	167.38	4.96	77.87	11.68	8.68	1.9	0.86	80.03	1.86	1
Finland	178.33	188.16	155.94	515.18	95.01	6,256.17	321.98	10,467.72	77.78	120.01	122.3	193.18	134.97	131.57	43.44	69.08	10.86	12.82	171.04	43.77	208.79
France	157.78	157.44	140.95	475.07	74.86	4,208.12	329.39	7,016.36	82.52	114.58	134.02	1,704.62	120.6	108.81	45.09	81.63	7.87	9.68	183.91	38.46	134.3
Georgia	129.2	124.57	130.76	386.74	64.73	1,223.27	236.46	2,660.71	54.02	110.25	110.22	74.96	62.27	61.48	29.75	17.48	7.24	2.59	125.9	7.94	47.28
Germany	171.45	182.06	146.26	484.82	84.71	4,761.93	324.27	6,751.83	78.54	110.69	104.56	24,235.65	128.7	137.86	43.46	97.94	13.12	9.38	179.56	40.5	159.81
Ghana	109.72	108.43	91.95	363.05	24.06	35.51	200.5	413.56	51.81	54.47	72.71	139.36	13.69	92.76	7.08	8.85	2.09	1.45	100.59	2.46	6.71
Greece	121.45	100.72	124.45	438.84	123.89	2,943.89	287.31	4,699.29	75.9	113.66	88.25	228.72	86.83	53.44	22.62	38.9	19.58	2.02	118.06	20.59	47.45
Guatemala	82.01	73.91	96.24	363.86	25.64	20.31	180.27	652.92	51.85	83.77	75.3	140.05	23.33	80.53	15.19	3.18	2.36	1.65	130.49	4.64	8.17
Guinea	77.65	36.94	103.57	291.09	21.33	107.94	99.78	284.34	46.08	29.22	144.3	86.94	8.95	107.87	12.26	1.85	1.55	0.04	119.28	1.3	1.23
Honduras	86.76	61.51	91.95	373.61	27.79	20.86	153.55	842.92	50.32	105.99	76.3	85.28	27.21	76.7	12.98	2.26	4.47	3.36	127.54	3.07	4.42
Hong Kong, China	182.59	188.17	138	504.39	83.09	3,193.03	323.97	4,495.85	80.03	211.42	755.33	336.42	94.84	125.86	74.03	41.18	16.62	20.72	160.25	41.84	193.02
Hungary	139.76	121.53	120.32	450.9	59.16	2,743.51	262.82	2,904.35	68.86	80.2	77.41	219.39	95.22	83.64	54.72	33.72	17.35	16.21	133.87	13.58	84.43
Iceland	170.28	168.89	145.33	462.87	73.75	5,643.32	296.32	48,089.94	74.77	112.76	132.86	68.43	112.93	106.21	23.48	39.68	11.38	9.58	178.93	89.83	237.4
India	107.78	85.94	102.56	338.17	45.87	308.56	230.18	995.44	33.58	96.55	132.51	5,800.41	39.73	109.46	22.86	39.87	4.49	8.61	118.2	2.83	10.67
Indonesia	111.66	100.96	125.54	378.57	43.4	109.29	197.6	855.57	49.68	80.62	102.89	1,990.46	43.95	102.45	27.6	11.08	3.27	1.52	126.26	2.41	7.7
Iran (Islamic Republic of)	95.96	47.6	87.06	403.64	87.52	634.86	219.23	3,093.38	55.6	87.4	70.49	970.21	53.92	69.78	12.02	32.61	3.8	0.53	199.12	2.6	14.44
Ireland	163.11	168.96	147.35	483.26	84.48	4,039.09	308.04	5,489.42	94.69	85.06	108.31	266.37	109.89	121.13	78.68	37.31	10.12	43.87	161.39	36.79	124.71
Israel	151.68	156.77	139.01	448.68	68.51	7,655.96	296.3	6,666.99	76.06	98.04	125.25	247.89	117.83	134.31	59.76	60.07	21.74	16.79	148.84	28.61	230.53
Italy	129.3	115.13	140.45	461.75	69.96	2,214.29	303.71	4,182.65	81.23	95.63	87.77	1,382.86	96.58	113.06	30.95	70.74	32.02	4.25	140.08	22.14	78.09

Jamaica	127.26	91.27	140.78	392.14	37.42	522.86	141.53	1,267.64	60.15	66.8	83.57	80.97	47.81	85.98	16.73	7.39	2.86	2.16	161.31	7.27	9
Japan	148	163.9	150.8	499.92	77.03	5,012.06	331.92	7,297.84	76.39	169.48	141.69	3,181.47	135.76	132.55	53.54	108.92	9.76	11.13	173.21	47.49	43.3
Jordan	111.76	97.78	96.58	382.37	50.8	565.35	199.03	1,797.94	61.29	83.34	102.98	116.08	50.11	88.26	15.18	16.35	4.51	0.29	121.08	4.87	35.51
Kazakhstan	114.37	82.99	135.39	396.42	62.12	657.18	280.31	5,118.48	51.76	70.26	104.32	330.45	79.56	67.51	29.47	6.29	2.76	2.74	101.07	14.56	18.97
Kenya	96.54	75.98	117.71	372.62	21.8	208.83	159.35	202.72	45.56	88.98	102.88	164.98	18.22	103.67	11.83	15.71	4.19	3.08	131.42	5.52	3.33
Kuwait	103.42	99.32	101.63	421.7	48.61	454.06	271.04	14,728.23	60.39	102.05	132.4	215.8	37.16	80.66	20.91	11.54	3.82	7.7	119.96	8.44	22.25
Kyrgyzstan	80.26	60.68	118.39	392.47	52.93	630.63	202.29	1,875.05	51.95	72.29	99.72	61.33	31.76	51.11	19.57	8.12	2.02	2.11	74.86	2.54	8.28
Latvia	145.57	144.51	129.42	463.34	89.32	1,646.11	262.02	3,328.31	69.67	108.1	119.49	93.03	79.01	78.42	18.6	14.79	16.97	6.21	152.46	36.7	124.29
Lebanon	76.58	60.21	91.1	351.57	51.6	767.76	194.68	2,669.12	60.81	110.3	74.83	116.86	54.07	80.04	16.84	68.69	6.98	3.59	85.23	8.28	35.18
Lithuania	152.45	146.14	117.96	449.23	76.42	2,775.37	277.4	1,127.58	74	82.74	122.29	115.53	98.67	85.45	22.62	20.29	12.51	3.93	151.58	23.56	149.68
Luxembourg	179.1	182.95	113	455.78	49.33	4,340.13	333.52	1,321.08	80.3	92.49	125.39	96.35	113.58	127.86	57.71	28.95	9.8	53.4	189.38	59.4	242.78
Madagascar	75.52	52.19	102.98	316.35	21.51	28	85.67	436.48	34.23	39.9	91.3	76.26	6.49	68.24	9.39	5.34	2.51	2.49	111.17	1.84	1.74
Malawi	89.86	64.16	93.04	347.39	14.5	44.48	75.06	322.77	47.39	73.48	141.85	65.95	5.97	59.32	12.8	11.22	1.64	1.6	73.6	1.45	1.63
Malaysia	147.42	124.45	126.3	397.73	60.63	2,244.44	291.99	4,313.35	59.74	145.95	167.03	616.8	87.29	128.55	26.57	19.14	10.46	16.09	140.93	9.95	20.46
Mali	69.99	56.91	107.44	325.37	17.98	30.29	92.73	371.17	43.35	42.85	119.36	78.78	4.31	82.13	25.04	4.76	1.52	3.99	86.08	0.91	6.64
Malta	153.04	150.86	103.57	444.08	56.1	1,938.49	310.86	1,684.00	89.38	86.78	98.76	77.81	82.35	81.41	59.42	16.33	15.56	9.39	191.43	36.52	135.04
Mauritius	155.56	163.1	137.58	403.71	49.42	166.54	243.93	2,078.12	60.72	126.18	122.45	74.29	30.74	78.32	14.16	5.25	10.23	1.63	116.85	10.18	23.92
Mexico	104.97	86.1	131.94	401.48	49.45	305.63	268.07	2,237.72	60.81	92.78	92.59	1,479.08	44.53	87.31	23.45	24.83	3.77	6.2	134.04	6.91	9.93
Mongolia	114.76	78.35	97.92	415.32	68.39	1,293.62	204.52	1,604.32	55.35	103.64	107.48	77.86	41.82	51.12	19.01	10.03	4.48	0.47	207.36	23.09	17.09
Montenegro	122.51	98.29	128.57	401.65	66.62	653.87	251.29	4,331.86	61.62	99.39	124.21	59.62	75.23	75.35	19.28	11.89	7.28	0.82	124.39	24.79	112.96
Morocco	107.01	84.35	122.76	406.21	41.49	981.88	229.87	786.72	65.14	81.28	105.7	230.76	40.58	68.94	8.98	12.92	4.32	3.1	134.57	2.19	7.63
Mozambique	81.23	58.49	96.41	336.71	12.22	38.26	113.42	570.69	42.96	37.86	136.74	68.69	4.6	69.96	20.33	6.01	3.19	0.55	93.75	2.11	1.49
Namibia	123.87	90.69	89.34	369.93	29.15	131.4	149.04	514.66	59.54	92.69	79.37	65.98	33.12	72.89	11.71	7.11	1.88	0.41	155.27	5.57	14.19
Nepal	88.5	62.19	110.81	361.95	23.54	230.19	197.31	148.27	31.29	97.98	174.37	107.45	27.04	62.58	16.99	13.12	2.32	3.44	94.92	2.05	8.33
Netherlands	177.94	191.78	150.38	487.52	79.78	4,722.54	334.92	5,858.75	75.93	118.87	139.39	603.23	118.69	139	70.39	78.82	11.74	40.02	175.01	38.42	223.75
New Zealand	181.14	191.41	144.66	485.35	90.2	3,797.53	332.75	7,661.38	65.44	190.38	107.75	168.29	95.38	108.37	28.87	47.75	10.98	1.42	174.78	41.99	138.49
Nicaragua	83.85	58.23	101.8	379.22	35.4	223.41	132.4	660.62	54.06	55.52	72.57	70.19	31.16	56.03	13.31	3.39	2.93	2.34	101.78	1.94	10.29
Niger	74.26	56.84	112.07	323.56	14.68	148.59	79.19	26.33	36.36	14.29	89.31	69.97	11.14	68.28	14.59	4.28	2.23	3.42	82.2	1.46	1.98
Nigeria	67.76	43.88	95.48	362.48	22.42	185.57	135.25	161.29	52.83	72.85	82.57	709.01	39.01	60.17	11.33	9.77	1.6	0.69	101.65	8	1.98
North Macedonia	117.38	98.04	47.83	354.37	49.2	667.59	244.85	2,326.01	65.79	80.69	113.98	69.7	63.22	59.75	19.41	10.72	12.87	3.32	111.4	11.34	57.02

Norway	184.43	187.11	151.31	482.93	39.16	5,955.66	329.75	23,906.86	78.63	152.63	119.89	278.14	121.21	115.45	33.29	47.76	9.48	4.78	151.64	25.62	190.04
Oman	129.2	114.8	113.84	462.7	69.74	231.03	278.87	6,605.13	50.75	83.22	89.36	164.18	62.1	99.75	4.45	6.8	4.23	2.13	129.04	5.04	19.7
Pakistan	76.39	64.12	119.4	365.33	27.69	291.47	141.67	517.07	39.6	45.91	98.06	687.77	30.42	84.16	15.57	15.81	2.77	1.84	108.38	0.61	6.03
Panama	117.67	104.48	57.85	413.7	52.7	38.87	227.11	2,345.72	67.91	125.22	82.01	122.56	40.82	73.45	8.4	12.16	2.86	3.13	131.79	13.44	71.35
Paraguay	88.21	76.44	100.03	404.18	44.58	111.91	180.91	8,075.71	54.3	60.38	92.16	108.95	24.99	55.51	18.55	4.22	3.71	0.58	194.72	5.99	7.1
Peru	105.94	88.44	107.86	379.45	74.4	599.82	239.99	1,415.29	65.69	88.51	99.05	312.26	62.99	64.37	18.8	11.4	5	0.44	121.04	6.95	11.44
Philippines	102.36	87.98	107.1	411.67	49.34	189.37	252.36	782.84	60.49	40.44	113.35	589.6	64.65	96.73	50.16	12.04	5.06	17.17	132.02	3.14	6.46
Poland	139.18	128.65	134.21	477.95	71.83	2,807.05	300.02	3,815.06	64.79	94.9	93.44	721.53	89.03	73.82	36.2	41.27	9.67	5.44	122.12	10.27	66.88
Portugal	160.69	145.13	143.9	474.67	73.97	4,048.58	311.84	4,812.84	74.83	112.9	91.55	239.66	91	98.95	27.52	43.45	19.46	3.28	180.54	26.58	71.9
Qatar	134.92	125.12	105.92	383.56	45.26	562.37	275.99	14,035.31	65.42	88.61	106.9	252.98	27.52	116.9	17.31	8.03	3.68	3.58	120.17	22.08	13.51
Republic of Korea	154.88	156.08	150.55	497.99	98.2	7,031.13	345.7	9,328.34	61	158.73	140.94	1,252.64	145.85	115.08	57.86	118.09	7.13	12.8	188.65	35.78	63.79
Republic of Moldova	93.64	84.67	126.05	408.78	51.13	662.65	266	1,411.56	49.07	71.82	99.14	68.03	52.07	51.83	9.48	12.13	5.38	3.34	159.18	4.96	36.46
Romania	110.01	111.07	121.08	416.28	61.71	815	248.72	1,978.32	73.84	78.41	77.34	336.01	75.37	66.77	23.28	19.11	21.58	5.46	117.06	6.63	39.66
Russian Federation	103.81	60.75	127.65	466.1	90.12	2,690.70	296.79	6,435.23	59.38	98.66	97.53	2,367.07	97.58	82.11	35.7	39.45	3.84	3.51	128.68	6.13	44.52
Rwanda	123.68	98.17	125.12	413.96	16.88	11.58	179.42	407.03	42.86	86.78	145.28	67.03	14.17	81	10.61	5.36	2.28	0.94	104.81	2.74	1.9
Saudi Arabia	104.59	98.66	67.44	451.71	73.62	2,610.02	263.9	9,109.75	55.94	74.37	126.23	1,089.03	70.27	99.31	27.43	19.54	2.37	1.94	116.26	14.54	9.24
Senegal	106.62	84.33	113	300.86	24.72	503.36	143.61	256.07	49.38	44.8	45.23	96.94	9.18	72.57	5.26	7.48	2.41	3.64	115.99	1.38	2.3
Serbia	122.03	84.07	129.16	420.14	74.87	1,905.78	269.07	4,677.06	62.44	80.3	77.59	116.37	49.97	74.9	15.04	28.92	20.55	4.21	112.96	9.89	57.73
Singapore	194.8	189.6	145.25	518.42	104.61	6,265.69	329.39	7,899.48	69.07	153.21	220.01	370.52	125.8	126.75	58.9	43.21	10.39	20.16	148.58	30.42	89.16
Slovakia	144.99	129.7	125.37	440.16	56.49	2,570.43	273.48	4,153.61	74.91	96.98	71.49	166.4	85.81	78.94	22.57	22.47	18.63	7.37	146.75	27.21	55.21
Slovenia	155.17	130.96	148.7	484.85	82.19	4,147.56	282.88	6,608.00	70.06	67.27	87.02	102.35	154.34	134.87	43.98	43.49	19.34	3.73	174.36	34.97	124.52
South Africa	98.09	95.07	114.26	408.16	31.84	506.44	243.84	3,842.48	43.77	157.29	265.56	497.47	65.49	100.6	16.84	30.32	9.4	2.71	114.22	7.05	14.16
Spain	145.19	143.13	139.77	468.23	92.11	2,740.48	321.66	5,042.80	82.51	125.02	117.89	1,093.69	95.56	87.56	28.59	61.42	15.6	6.73	156.69	23.67	96.97
Sri Lanka	112.24	110.77	111.99	381.48	32.3	100.82	185.33	597.69	67.36	64.33	91.36	220.68	55.71	76.2	19	9.9	3.32	3.14	106.6	3.65	7.93
Sweden	177.65	187.34	146.68	477.82	73.41	6,780.02	328.71	13,583.98	83.94	141.77	126.17	361.67	163.08	131.32	52.66	80.62	10.63	12.8	176.33	52.54	231.52
Switzerland	186.18	187.31	127.23	481.58	72.18	4,971.45	317.85	6,101.43	93.56	178.59	205.05	366.67	143.92	146.43	45.49	93.31	17.48	16.27	188.7	71.95	194.51
Tajikistan	74.16	33.45	102.39	327.8	41.24	272.77	129.95	1,687.56	48.03	39.61	211.47	69.93	21.15	77.9	16.68	2.35	1.6	2.49	87.72	2.94	4.42
Thailand	123.19	106.77	142.55	408.22	60.58	1,175.58	223.24	2,415.34	51.01	161.55	145.94	792.94	86.88	90.79	43.06	20	7.27	8.04	127.79	5.49	13.84
Togo	82.59	52.66	114.01	342.8	23.89	35.33	152.05	44.55	39.49	53.04	87.12	66.46	12.52	68.24	7.27	2.96	2.23	6.75	88.23	2.59	1.92
Trinidad and Tobago	123.58	92.54	115.44	408.63	62.05	1,459.28	230.52	6,655.25	61.42	78.38	110.8	84.73	65.67	65.7	17.9	4.87	4.34	0.38	101.96	14.54	21

Tunisia	107.4	86.53	121.67	378.19	58.88	1,798.54	240.59	1,498.06	63.37	102.8	83.38	138.91	41.6	67.83	7.15	23.16	7.34	2.75	114.8	2.48	7
Turkey	110.59	93.73	108.53	411.46	98.9	1,335.76	269.82	3,239.93	56.12	108.35	94.22	1,345.18	71.96	73.3	39.59	30.05	4.15	0.91	150.07	6.4	29.39
Uganda	90.15	75.75	94.47	365.08	21.03	24.44	149.25	370.85	43.33	55.6	156.4	117.18	16.12	81.95	7.2	10.85	2.15	0.46	90.57	2.81	2.34
Ukraine	79.49	64.76	103.4	402.41	86.25	1,044.10	213.06	3,106.29	49.43	84.02	85.15	255.01	86.98	73.08	21.98	30.5	4.78	4.32	164.32	3.82	52.93
United Arab Emirates	157.11	135.1	121.08	449.33	85.01	2,295.08	326.32	11,947.31	59.15	111.55	119.22	465.44	94.05	136.06	44.85	10.62	6.7	5.13	127.58	20.45	29.04
United Kingdom	155.75	176.26	147.27	483.45	74.91	4,170.89	341.84	4,336.09	84.36	159.38	137.59	1,740.53	119.73	131.16	33.15	99.8	17.12	8.92	174.21	45.49	187.63
United Republic of Tanzania	87.43	63.04	94.05	318.48	17.06	17.22	141	128.94	48.52	58.3	159.62	151.31	3.75	88.68	9.08	8.85	2.18	0.21	103.16	2.72	1.9
United States of America	164.18	172.47	153.33	467.97	86.82	4,077.07	330.04	11,112.91	68.45	217.03	167.71	11,392.59	143.1	147.01	53.96	100.65	5.98	8.35	158.13	67.16	128.76
Uruguay	137.44	126.7	120.24	414.39	62.04	621.89	300.7	3,294.11	67.73	63.81	101.52	99.31	33.03	65.29	5.32	14.19	12.52	3.66	135.4	16.49	77.04
Viet Nam	126.1	88.11	100.79	479.03	39.38	650.42	211.83	1,561.10	47.45	156.09	91.98	443.66	62.07	75.41	29.18	13.34	8.41	12.83	142.16	2.8	41.03
Yemen	1	15.99	78.22	358.51	24.13	130.33	68.41	163.99	53.67	4.34	0.67	84.82	6.83	46.03	7.11	3.43	0.46	0.67	60.92	1.62	2.5
Zambia	96.16	88.65	92.2	364.51	17.77	37.8	131.97	622.08	48.84	77.59	74.55	98.56	9.61	65.99	8.83	6.69	1.94	0.31	81.45	3.21	3.94
Zimbabwe	64.08	21.58	77.3	323.89	28.97	81.46	109.86	375.25	41.19	66.26	130.47	70.86	12.8	50.06	8.61	9.47	5.57	0.29	67.48	4.08	2.78

Source: own elaboration.

Appendix 3: Descriptive analysis for the four clusters after correction.

	Indicator	Mean				Std				Min				Max			
		Specials	Advanced	Intermediates	Primitives	Specials	Advanced	Intermediates	Primitives	Specials	Advanced	Intermediates	Primitives	Specials	Advanced	Intermediates	Primitives
1	Political and operational stability	89.96	80.65	68.07	57.89	6.51	5.59	9.38	13.01	75.4	68.4	43.9	0	100	87.7	93	84.2
2	Government effectiveness	85.56	66.97	47.81	31.13	9.89	8.13	9.24	11.18	58.7	52.9	34.1	0	100	80.4	74.3	58.4
3	Regulatory quality	87.38	64.90	44.97	26.87	8.64	10.89	12.25	11.43	71.6	38	9.8	0	100	86	77.8	54.4
4	Rule of law	90.61	64.61	43.78	28.83	7.35	11.92	11.76	12.63	73.4	39.4	24.6	0	100	80.3	73.1	60
5	Cost of redundancy dismissal	13.71	14.61	19.68	21.72	6.60	5.78	12.60	11.72	8	8	8	8	27.4	27.4	73.6	58.5
6	Ease of starting a business	92.93	89.08	83.82	81.26	4.29	4.56	14.36	9.73	83.6	82	29.1	52.8	100	95.3	99.3	97.4
7	Ease of resolving insolvency	78.17	64.82	49.64	40.19	11.97	13.41	15.26	10.33	45.5	38.1	0	24.4	93	83.7	76.6	69.8
8	Expenditure on education	5.34	4.80	4.26	4.31	1.51	0.62	1.28	1.71	2.9	4.1	0.7	1.1	7.7	6.4	7.4	9.6
9	Government funding/pupil, secondary	23.72	23.17	20.20	18.38	7.96	5.19	8.66	8.52	16.5	17.1	5.8	5.1	52.8	39.4	52.8	44
10	School life expectancy	17.21	15.73	14.51	10.97	1.57	1.41	1.39	2.01	14.2	13.5	11.3	6.5	19.3	17.9	17.7	15.4
11	PISA scales in reading, maths, & science	506.19	477.86	421.49	365.51	19.83	28.88	34.58	36.30	471.7	412.7	336	286.892	551.6	524.3	491.8	419.94
12	Pupil-teacher ratio, secondary	12.30	10.06	13.73	22.67	2.51	1.73	5.69	7.44	8.7	7.2	6.7	10.4	19.4	13.3	28.5	40.4
13	Tertiary enrolment	68.67	67.23	52.54	17.10	19.41	20.03	20.03	12.77	19.6	41.9	16.4	0.8	93.8	126.4	103.7	47.7
14	Graduates in science & engineering	24.36	23.55	24.25	18.19	5.93	3.96	7.43	4.64	14.1	15.9	11.6	9	36	32.1	44.8	31.1
15	Tertiary inbound mobility	12.02	8.63	4.25	2.59	11.02	10.70	5.80	2.43	1.9	0.4	0.1	0	47	48.6	35.3	10.7
16	Researchers, FTE/mn pop	5506.22	2738.72	946.27	192.50	1372.04	921.41	765.54	205.22	3411.7	1174.4	39.1	12.3	8250.5	4467.8	3154.85	820.8
17	Gross expenditure on R&D	2.47	1.17	0.45	0.30	1.11	0.46	0.33	0.19	0.8	0.5	0.1	0	4.6	2.1	1.34	0.8
18	Global R&D companies, avg. exp.	72.28	34.96	10.63	0.00	25.29	31.80	21.62	0.00	0	0	0	0	100	91.7	73.9	0
19	QS university ranking	59.68	25.88	16.20	1.78	27.35	20.60	15.96	5.57	0	0	0	0	99	82.5	47.3	25.7
20	ICT access	86.11	76.89	66.36	37.89	5.54	6.62	10.73	9.73	73.9	60	38.5	20.8	94.2	91.3	82.5	55.6
21	ICT use	82.34	70.27	54.24	23.99	4.54	6.52	10.57	12.53	75.8	61.5	20.8	6.1	89.7	81.3	73.2	63.8

22	Government’s online service	92.06	83.03	72.96	46.85	7.09	9.50	12.54	17.65	72.9	65.3	43.1	9.7	98.6	95.1	95.1	78.5
23	E-participation	93.38	85.17	74.00	45.91	7.97	10.19	13.62	18.59	68.5	61.8	43.3	11.8	100	98.9	97.2	80.3
24	Electricity output	12176.35	5710.68	4534.25	856.03	12438.88	2915.21	4325.79	1476.15	1480	1280.8	878.8	25.4	56585.3	13980.2	19937.1	9475.6
25	Logistics performance	81.91	60.22	37.52	22.09	12.40	16.14	12.93	10.43	54.3	34.9	14.5	0	100	88.5	66	47.6
26	Gross capital formation	23.60	22.35	24.89	24.72	3.69	6.48	5.63	10.66	17.2	12.6	15.8	5.9	31.2	44.2	39.1	51.8
27	GDP/unit of energy use	12.41	10.58	9.45	8.09	6.07	4.27	3.91	3.98	2.7	6.6	2.2	2.4	27	25.6	18.9	21.3
28	Environmental performance	76.15	68.60	58.52	47.34	7.20	7.68	7.44	8.41	62.3	50.7	30.6	27.4	87.8	80.9	67.9	61.2
29	ISO 14001 environmental certificates	3.54	6.87	1.94	0.48	2.14	3.38	2.82	0.72	0.3	1.8	0.1	0	9.3	13.5	11.7	4.6
30	Ease of getting credit	63.89	62.50	63.14	50.38	18.67	13.53	19.21	26.06	15	35	5	0	100	85	100	95
31	Domestic credit to private sector	126.22	82.49	59.71	29.92	43.64	42.91	29.87	18.83	44.4	33.4	16.1	5.6	203.8	199.1	147.7	86.7
32	Microfinance gross loans	0.18	0.20	0.83	1.86	0.03	0.10	2.87	4.50	0.135	0	0	0	0.269	0.402	18.5	28.2
33	Ease of protecting minority investors	67.47	63.61	63.19	49.27	10.03	7.70	13.07	12.05	48.3	50	28.3	28.3	81.7	81.7	85.3	78.3
34	Market capitalization	156.35	46.76	53.03	66.91	241.15	33.28	45.55	74.62	34.4	5.1	4.6	-111.017	1099.6	131.7	302.1	352.27
35	Venture capital deals	0.16	0.03	0.02	0.01	0.12	0.05	0.04	0.02	0	0	0	0	0.4	0.1	0.2	0.1
36	Applied tariff rate	1.86	2.20	3.78	7.89	1.07	0.94	2.91	4.00	0	1.8	0	0.9	5.1	4.8	15.2	17.8
37	Intensity of local competition	76.43	72.92	67.58	62.82	6.74	5.22	7.50	6.79	61.7	59.3	46.4	45.6	87.2	80.4	86.2	75.9
38	Domestic market scale	4602.23	1919.04	906.21	187.67	10884.74	5877.37	1789.18	322.93	19.3	20.8	11.8	8.2	43791	25313.3	10401.4	1297
39	Knowledge-intensive employment	46.48	33.90	24.40	10.76	7.71	7.38	7.83	7.00	25.2	18.4	6.9	1.1	59.1	45.5	44.3	30.3
40	Firms offering formal training	40.57	38.98	36.13	33.49	10.12	13.61	16.30	14.59	18.6	15.8	3.4	10	70.3	79.2	65.1	73.7
41	GERD performed by business	1.70	0.71	0.22	0.05	0.99	0.41	0.21	0.08	0.4	0.1	0	0	3.9	1.7	0.853	0.35
42	GERD financed by business	54.29	50.59	28.76	8.86	11.91	17.41	15.45	9.26	34.7	21.6	0.3	0.1	78.3	96.2	75.2	40.2
43	Females employed	21.39	17.79	12.73	3.85	4.43	5.74	5.94	3.31	13.2	8.8	1.6	0.2	28.4	27.9	32.6	10.8
44	University/industry research collaboration	68.95	47.95	40.95	35.81	7.80	12.12	9.23	9.30	54.6	25.6	26.6	19.5	80.9	74.1	64.8	67.2
45	State of cluster development	66.66	53.94	45.76	40.48	7.38	15.88	9.37	7.41	52.6	32.3	28.2	26.6	79.5	96.1	65.4	55.4

46	GERD financed by abroad	11.66	13.51	6.30	18.16	11.59	8.90	5.62	15.56	0.6	0.6	0.1	0	49.8	34.2	24.2	52.4
47	JV-strategic alliance deals	0.09	0.03	0.01	0.01	0.07	0.08	0.03	0.03	0	0	0	0	0.2	0.3	0.2	0.1
48	Patent families 2+ offices	5.55	0.86	0.07	0.02	3.91	1.23	0.08	0.08	1.1	0	0	0	14.4	5.3	0.3	0.5
49	Intellectual property payments	3.21	0.93	0.92	0.33	5.08	0.75	0.82	0.41	0.3	0.2	0	0	22.2	3.6	3	1.6
50	High-tech imports	12.41	10.43	7.71	7.32	10.39	6.31	3.93	3.59	1.9	4.6	1.5	2	49.8	26.4	23.2	21.8
51	ICT services imports	2.19	1.67	1.07	1.03	1.01	1.64	0.76	0.82	0.3	0.8	0	0	4.2	8.1	3.9	3.5
52	FDI net inflows	10.63	6.90	3.17	3.72	15.51	10.96	3.28	4.17	-9.3	0.8	-7.8	-0.8	45.3	43.8	11.6	24.3
53	Research talent, % in business enterprise	56.59	41.16	22.60	9.60	15.00	15.24	14.89	7.00	36.9	18.6	0.4	0.1	83.7	62.2	63.2	31.4
54	Patents by origin	14.52	5.62	1.65	0.38	19.05	12.22	1.90	0.92	0.7	0.1	0	0	78.2	53.7	9.3	6
55	PCT patents by origin	4.42	0.81	0.25	0.05	2.64	0.64	0.36	0.09	1.4	0.1	0	0	8.8	2.2	1.7	0.393
56	Utility models by origin	1.16	5.00	2.42	1.39	0.70	16.84	4.12	2.19	0.1	0	0	0	3.2	72.4	24.3	13.28
57	Scientific & technical articles	20.72	17.56	10.20	3.95	7.70	8.45	16.04	2.24	9.2	3.1	0.3	0.5	34.1	35.4	103	9.7
58	Citable documents H-index	54.02	25.81	12.67	5.49	26.53	18.55	10.08	3.75	9.1	5.4	0.6	0	100	69.2	38.9	15.5
59	Growth rate of PPP$ GDP/worker	0.94	1.87	1.35	1.49	0.98	1.82	2.38	2.65	-0.9	-0.2	-4	-11.1	2.8	7.1	6.2	5.5
60	New businesses	8.16	6.31	2.85	1.25	6.76	6.08	2.66	2.81	0.2	0.8	0.1	0	27.3	20.8	10.2	18.4
61	Computer software spending	0.67	0.34	0.22	0.10	0.63	0.19	0.13	0.13	0.2	0.1	0	0	3	0.7	0.5	0.4
62	ISO 9001 quality certificates	9.53	21.17	6.64	1.52	6.12	9.38	7.28	1.35	1.5	6	0.8	0	27.1	42.2	27	6.1
63	High- & medium-high-tech manufactures	0.41	0.32	0.22	0.10	0.22	0.17	0.13	0.06	0.1	0.1	0	0	0.8	0.6	0.5	0.216
64	Intellectual property receipts	2.64	0.44	0.20	0.06	1.88	0.66	0.34	0.13	0.1	0	0	-0.1	7	2.5	1.22	0.6
65	High-tech net exports	9.06	8.65	2.88	0.45	7.78	9.20	5.59	0.66	0.1	0.1	0	0	27.4	34.1	32.7	2.7
66	ICT services exports	4.36	2.69	2.05	1.49	5.44	3.10	2.15	1.38	0.4	0.5	0	0	22.7	14.6	10.4	5
67	FDI net outflows	11.59	5.18	1.00	0.49	16.72	11.18	1.54	1.11	-0.3	0.1	-2.7	-0.3	63.5	48.5	6.4	6.9
68	Trademarks by origin	62.96	75.34	51.69	32.49	28.51	50.94	46.89	40.73	11.8	8.7	3.3	2.8	102.9	238.7	229.1	220.4
69	Industrial designs by origin	5.63	7.23	3.36	1.15	6.89	6.50	4.93	1.38	0.8	0.1	0	0	29.7	26.3	18	5.5
70	ICTs & business model creation	78.76	67.17	58.84	51.44	4.03	7.09	6.76	9.31	72.8	52.7	43.2	24.3	84.7	78.6	72.1	68.9
71	ICTs & organizational model creation	75.03	61.53	52.45	44.42	5.51	8.60	6.90	9.12	64	44.6	39	21.7	83.7	79.3	65.4	60.6

72	Cultural & creative services exports,	1.25	1.06	0.58	0.18	1.02	1.28	0.76	0.32	0.1	0.1	0	0	4	5.9	4.2	1.5
73	National feature films	14.93	8.15	4.38	1.51	22.68	5.31	4.75	1.75	2.9	0.8	0.1	0.1	98.5	19.8	26.3	11.2
74	Entertainment & Media market	56.80	24.93	9.06	1.74	17.92	11.52	7.58	1.93	29.3	6.9	0.5	0	100	44.90	31.98	7.06
75	Printing & other media	1.38	2.54	1.38	1.34	0.96	4.99	0.90	0.64	0.3	0.5	0.2	0.2	4.8	22.4	4.1	4.2
76	Creative goods exports	2.47	3.87	1.23	0.21	2.32	3.78	2.16	0.30	0.1	0.2	0	0	9.9	11.9	9.6	1.3
77	Generic top-level domains	51.86	21.33	5.44	0.95	27.60	24.26	9.98	1.48	8.1	2.4	0.3	0	100	93.7	66.4	8.5
78	Country-code TLDs	44.97	21.06	5.83	0.52	33.88	13.24	15.13	1.05	2.4	3.3	0.2	0	100	48.5	100	6.7
79	Wikipedia edits	69.40	46.16	16.64	1.83	36.19	33.05	17.91	2.09	17.7	0.3	1	0	148.4	133.5	102.5	7.3
80	Mobile app creation	39.56	22.40	8.85	1.11	29.44	31.24	12.50	1.57	1.2	2.8	0	0	100	100	66.5	5.25

Source: own elaboration.

Appendix 4: The list of the GII indicators that have negligible mean differences across the four clusters.

	Indicators
In1	Political and operational stability
In3	Regulatory quality
In6	Policies for doing business
In8	Expenditure on education
In9	Government funding/pupil, secondary, % GDP/cap
In15	Tertiary inbound mobility, %
In24	Electricity output, GWh/mn pop.
In28	Environmental performance
In29	ISO 14001 environmental certificates/bn PPP$ GDP
In31	Domestic credit to private sector, % GDP
In34	Venture capital investors, deals/bn PPP$ GDP
In35	Venture capital recipients, deals/bn PPP$ GDP
In37	Applied tariff rate, weighted avg., %
In40	Knowledge-intensive employment, %
In42	GERD performed by business, % GDP
In44	Females employed w/advanced degrees, %
In45	University-industry R&D collaboration
In50	Intellectual property payments, % total trade
In51	High-tech imports, % total trade
Out1	Patents by origin/bn PPP$ GDP
Out2	PCT patents by origin/bn PPP$ GDP
Out4	Scientific and technical articles/bn PPP$ GDP
Out8	Software spending, % GDP
Out9	ISO 9001 quality certificates/bn PPP$ GDP
Out10	High-tech manufacturing, %
Out12	Production and export complexity
Out15	Intangible asset intensity, top 15, %
Out16	Trademarks by origin/bn PPP$ GDP
Out17	Global brand value, top 5,000, % GDP
Out18	Industrial designs by origin/bn PPP$ GDP
Out19	Cultural and creative services exports, % total trade
Out20	National feature films/mn pop. 15–69
Out23	Creative goods exports, % total trade
Out25	Country-code TLDs/th pop. 15–69
Out26	GitHub commit pushes received/mn pop. 15–69

Source: own elaboration.