May 2017

Facility Location Decision for Global Entrepreneurial Small-to-Medium Enterprises Using Similarity Coefficient-based Clustering Algorithms

Suhail H. Serbaya
University of Wisconsin-Milwaukee

FACILITY LOCATION DECISION FOR GLOBAL ENTREPRENEURIAL SMALL-TO-MEDIUM ENTERPRISES USING SIMILARITY COEFFICIENT-BASED CLUSTERING ALGORITHMS

by

Suhail Hasan Serbaya

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Engineering

at

The University of Wisconsin-Milwaukee

May 2017

ABSTRACT

FACILITY LOCATION DECISION FOR GLOBAL ENTREPRENEURIAL SMALL-TO-MEDIUM ENTERPRISES USING SIMILARITY COEFFICIENT-BASED CLUSTERING ALGORITHMS

by

Suhail Hasan Serbaya

The University of Wisconsin-Milwaukee, 2017
Under the Supervision of Associate Professor Hamid Seifoddini

Decisions on location selection are critical for the survival of small-to-medium entrepreneurial organizations from the time they are established until later stages of operation and expansion. The selection of location for small and medium entrepreneurial businesses requires a selection strategy that incorporates relevant factors, quantifies these factors and develops a methodology that analyzes data for better decision-making. In the era of globalization where borders have become easier to transcend, many small ventures tend to choose more attractive international markets as a potential location for their operations where they can obtain higher returns on their investment. Thus, significant changes in the location decision process of the small and medium entrepreneurial companies have received great attention in the literature about small firms with global orientation as a response to the international entrepreneurship phenomenon. Therefore, consideration should be given to factors and attributes that reinforce the appeal of the international market to new businesses. These factors and attributes will provide the decision maker with an effective methodology for data analysis that will provide a framework for decision-making in the selection of locations for the entrepreneurial organization.

In this research, the most frequent and critical attributes to select the best location for the entrepreneurial firms (globally) are extracted from relevant literature. Then, a similarity-based

cluster analysis approach is introduced to quantify these attributes based on the existing data of economic metrics, such as technological advancement, expenditures on education, expenditures on research and development, the quality of the labor force, unemployment rates, domestic competitiveness, etc. Subsequently, the resulting outcomes are used to identify groups of prospective sites that fit the needs of the entrepreneurial firm. Last, the validity of the adopted methodology will be tested via numerical examples.

Keywords: Entrepreneurship; Facility Location; Global Market; Location Decision; Small Ventures, SMEs

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

My deepest dedications must go to my mother Srrh and my father Hasan who passed away while I was in the middle of my studies. Their bodies might be gone but their kind and caring words will always be alive in me and have led me all the way. Their "never let go of your dreams" spiritual words have helped me through all the hard times.

I also express special thanks to my beloved sister Sahar and two brothers, Sameer and Saud. Their encouragement and enthusiasm have continuously inspired me to complete this work.

Towards the fulfillment of this research, I had the privilege to work with a distinguished mentor and recognized researcher, my advisor, Professor Hamid Seifoddini. His enlightenment and profound notes provided substantial input that widened my views on how to better carry out various components of the research. I was also fortunate to work closely with an outstanding teacher and researcher, Wilkistar Otieno, Ph.D. Her sincere advices and professional assistance added significant value to this research

The credit too goes to Mrs. Elizabeth Warras (Betty). Betty's devotion and cooperation have helped me in overcoming difficulties in all the downtimes I encountered within this work in particular and all aspects of my Ph.D. program.

Finally, sincere and great appreciation for the love of my life, my wife Hala. There are not enough words that can proclaim her unconditional love and overwhelming support. It is always enough to know that she will be there for me and by my side to keep me going and believe that my ambitions can go beyond the sky, not only in this work, but in every step in life.

CHAPTER ONE

Introduction

Making a decision on the facility location is a crucial factor for all types of organizations. Although it occurs infrequently, it is one of the most costly decisions that a company can encounter. Thus, business executives are required to conduct extensive research to properly identify the most suitable location for establishing their facility in order to guarantee a higher success rate for the business and to insure more efficient utilization of invested capital.

The facility location is an important decision because it requires large investments that are not recovered. Decisions on facility location have a great impact on the competitive capacity of the organization and other important aspects of the business such as operations, business development, human resource, finance, etc.

Furthermore, the facility location decision has a great influence on additional costs of the business (e.g., land, labor, raw materials, transportation and distribution costs) and on the firm's income. For example, proximity to the needed resources could greatly reduce the cost of shipping and transporting the goods to target markets.

Identifying the best location is even more important for small and medium businesses due to their tight budgets and limited resources. The decision of choosing a best location for small and medium enterprises has more influence on their business operations than on their large businesses counterparts, which might operate in multiple locations. Small and medium businesses might have a single location, making the decision to select another location a crucial factor in their long-term success.

1.1 Entrepreneurship definition and its importance to the economy

There are several definitions to describe the concept of entrepreneurship. One comprehensive definition is the process of creating something different with value by devoting the necessary time and effort; assuming the accompanying financial, psychological, and social risks, and receiving the resulting rewards of monetary and personal satisfaction and independence (Hisrich, Peters & Shepherd, 2007). Another significant definition of entrepreneurship is a scholarly examination of how, by whom, and with what effects opportunities to create future goods and services are discovered, evaluated, and exploited (Shane & Venkataraman, 2000).

More broadly, entrepreneurship can be defined as the process of gathering and allocating all necessary resources including financial, creative, managerial, and technological resources, to be successful in starting up and running a small enterprise that is based on a novel idea to fulfill the needs of prospective consumers for specific products or services. Successful entrepreneurship relies to a great extent on the dedication, talent, and creativity that the entrepreneur must possess. These distinguishing traits should be combined with innovative ideas, energy, and a clear vision in order for the entrepreneur to create the new venture. However, starting up a new venture requires more than just having a good business idea. Developing an effective business plan and forming a team of talented, experienced individuals to help manage the new business's operations are also critical to exploit the identified opportunity for profit.

1.1.1 Characteristics of entrepreneurship

Various significant features characterize the broad concept of entrepreneurship, including:

- An economic and dynamic activity

Entrepreneurship involves the creation and operation of a small enterprise in which the focus is on optimizing the exploitation of available resources to create value and wealth. Therefore, it is an economic activity.

On the other hand, the act of entrepreneurship is often performed in a business environment that is characterized by uncertainty. Thus entrepreneurship is considered as a dynamic activity.

- Integrated with innovation

  Entrepreneurship is all about searching for new business ideas including exploring more efficient approaches to carry on the related business operations. The entrepreneur continuously seeks innovation and optimization of performance in all aspects of the organization.

- Generates profit

  The added value through entrepreneurial activities is usually rewarded with obtaining profit that is an important motivation for entrepreneurs to translate their business ideas into a realistic venture.

- Involves risk-taking

  Start-up ventures based on innovative new ideas convey a lot of uncertainty. Therefore, entrepreneurship is typically associated with the capability of the entrepreneur to tolerate risk and pursue the new business venture.

1.1.2 Importance of entrepreneurship

Entrepreneurship brings important benefits to the economy. Some of these significant benefits are:

- Creation of new businesses and subsequently producing new employment opportunities

- Considerable contribution to the national income
- Creation of social change
- Development of the community

1.2 Statement of the problem

Promoting entrepreneurial practices is of great value for most countries, and specifically for developing countries, entrepreneurial activities are a major tool to enhance their economies. There are many attributes and factors, both tangible and intangible that require extensive measurement and evaluation in order to assist governments in their quest to meet the ongoing desires of economic and social prosperity. It is also important for the founder of the new firm/the entrepreneur to assess drivers of location-fit decision when either planning to establish their new venture, to explore the possibility for extension or to go global. Furthermore, the decision-making about location, in most of the cases, is a highly complex process.

The problem of choosing the best location of the facility has been and continues to be a focus of interest for many entrepreneurial scholars and researchers. In this realm they introduce algorithms and simpler software tools and packages to facilitate the location decision process for decision makers who are involved in the entrepreneurial activities. In order to make these algorithms efficient and to generate valid outputs, involved decision makers have to: (1) determine the type of the facility function they desire to best fit in a location, and (2) provide the most relevant combination of decision-making factors. Depending on the facility function type and the decision factors, the necessary data that formulate the inputs for the algorithm could be ready after verifying their accuracy and error-free status.

The relevant list of decision factors should be given great consideration by the involved decision makers since they constitute the pillar of function of all location decision algorithms while the absence of a well-prepared decision factors list could greatly impede the ability to identify the best solution.

The problem of facility location differs among firms. Therefore, the core industry of the facility, the produced goods, the type of targeted consumers and related variables are important considerations when dealing with the location decision.

The solution obtained for the facility location problem within one type of facility depends on related decision factors that cannot readily be applied to other types. However, there are multiple decision-making factors that are common for all types of firms. These common factors have been the focus of attention for many researchers who have offered various lists of these factors.

Locating international facilities is one aspect of the facility location problem that has attracted significant attention from scholars and researchers in recent years. Consistent with the growing trends of globalization and open international markets, researchers have provided the decision makers with practical forecasting tools to improve their capabilities in determining better options for locating their facilities in different countries. Many decision location factors in a specific country are fixed, but those similar factors differ from country to country and thus they should be studied and assessed to avoid irrelevant or unsound decisions.

Traditionally, the location decision for a facility was mainly linked to its proximity to required natural resources. Recent orientation to decide a best location for a facility considers a broader combination of factors such as rapid advancement in technology, improvements in

production methodologies, etc. The location decision also is affected by the more turbulent political world of today and natural or economic global disasters.

On a continuous basis, governments all over the world strive to define multiple means to assure the development of their region/nation both economically and socially. One major option they consider is flourishing productive entrepreneurial activities, as they are a principal source of economic growth and wealth creation. On the other hand, entrepreneurs and small venture founders seek all possible tools to reduce the related risk in establishing their new firms and maintaining their sustainability and growth in the context of supportive investment climates offered by regional and national governments. These reasons have stimulated both entities to pay more attention to the studies of international facility location decisions.

Furthermore, rapid changes in the global economy environment that in turn have a higher influence on local and regional economies have induced entrepreneurial organizations to explore more efficient ways to decide upon potential optimal international location for their activities. Many studies conducted by economists and entrepreneurship scholars have attempted to introduce possible forecasts. Their approaches vary from discussing entrepreneurial-attracting factors existing in specific geographical regions that contain several countries (attributes-based approach) to identifying the factors an individual country offers to attract entrepreneurial ventures (location-based approach).

These types of literature help to provide governments that constantly seek useful tools for their regions' or countries' prosperity via reinforcing the factors to encourage the entrepreneurial climate attractiveness in their specific economy. The literature also assists the founders of small firms in their location decision process to determine whether these reviewed regions or countries

have the requirements to be nominated as suitable locations for their entrepreneurial endeavors. Yet, this literature does not adequately convey enough information to comprise an efficient means to give the entrepreneurs a complete picture on all available alternatives so they can better decide what is the best location for their ventures.

Ranking the countries depending on their entrepreneurial attractiveness for small firms is considered a possible method to identify best-fit location for entrepreneurial ventures. Such rankings can be found in or inferred from several authenticated documents that are published by major entities such as the World Bank, Global Entrepreneurship Monitor (GEM), International Labor Organization (ILO), etc. However, these rankings could be misleading because they may not take into account the most influential location decision factors for entrepreneurs. On the other hand, a slight difference or error in a country's statistical data would result in assigning a specific country a lower rank than other countries, which deprives the decision maker of choosing a more suitable alternative.

In order to reduce the probability of misleading ranking, countries with convergent data could be classified and assigned into one group. Categorizing the countries in this form would leave the involved decision makers with more alternatives; they could identify a list of candidate countries to locate the firm instead of only nominating one country solely relying on its ranking. A further assessment among the group would then be carried out to determine the country that satisfies the specific requirements of the company.

Classifying countries on their similarities and dissimilarities can be carried out through various methods. One of the most efficient methods in data mining is clustering analysis, which also has the potential to accurately identify a specific framework in the studied data.

Furthermore, the preferred algorithm for categorizing the countries has to allow higher flexibility for the involved decision makers to define the measure of similarity depending on their needs. Hierarchical clustering can fulfill that purpose in addition to its capacity in testing a large amount of data in a short period of time.

In this context, this research addresses the problem of no available quantitative approach based on clustering algorithms to select the best location for entrepreneurial facilities while combining the most critical attractive factors to entrepreneurs.

1.3 Purpose of the research

The ultimate purpose of this thesis research is to create distinctive clusters that consist of homogenous groups of countries to promote the decision-making process of entrepreneurs who want to establish their new businesses internationally. The formed clusters also benefit the policy makers responsible for economical and social development by providing them with a comprehensive and efficient checklist to evaluate the status of their regions/countries' attractiveness to new entrepreneurial businesses compared with those countries that lie in other clusters.

Identifying and collecting the most critical attracting attributes to the entrepreneurial activities in order to prepare a comprehensive list of location decision-making factors is another major purpose of the research. This list is substantial for the process of creating clusters as well as determining what factors are missing for some regions or countries that could reinforce their attractiveness for entrepreneurs.

1.4 Objectives of the research

The main objectives of this research are:

- Identifying the most frequently cited attributes that attract entrepreneurial activities to a business location based on a relevant literature review.
- Applying the existing economic metrics such as technological advancement, expenditures on education, expenditures on research and development, the quality of the labor force, unemployment rates, and domestic business competitiveness, etc., for quantifying the attributes.
- Applying a similarity-based clustering algorithm to classify potential locations for entrepreneurship based on the most relevant attributes.
- Providing the decision makers in entrepreneurial firms with a flexible quantitative approach for selecting the best location for their entrepreneurial activities by allowing the users to include as many factors as necessary for particular applications.

1.5 Significance of the research

Defining the best-fit location for the entrepreneurial facilities through the application of similarity coefficient based clustering method offers the decision maker in the newly established company many advantages, including:

- Providing a highly flexible framework to facilitate the decision-making process of selecting the best location for entrepreneurial facilities.
- Quantifying the critical factors for entrepreneurial activities.
- Decreasing the reliance on surveys and questionnaires in which human judgment and opinion play a major role in the application of the existing methodologies.
- Elevating the ability to comprehensively compare large number of possible sites, an ability that also is lacking in the current location decision-making strategies.

- Applying similarity coefficient based clustering methods to identify groups of locations with similar characteristics, which have been applied successfully in the field of manufacturing, particularly cellular manufacturing, but have not been used previously for identifying potential locations for the entrepreneurial facilities.
- Providing the decision makers in charge with a convenient tool to choose the best-fit location for the entrepreneurial facilities/activities among multiple alternatives of locations that have similar output objectives. This method contrasts the previous approaches that proposed potential locations in the form of ranking only, in which even a small margin of error might result in losing a location's selection to another.
- Restricting the potential locations to accommodate the entrepreneurial facility, into a limited number of clusters that consist of similar countries instead of the far larger pool of individual countries to compare, evaluate and then choose the best alternative among them.
- Offering a unique classification of the studied locations into groups based on the strength level of the identified location decision-making factor(s).
- Allowing the entrepreneurs to customize the solution in accordance with their specific requirements and needs.
- The developed model is also applicable to the location decisions for starting new businesses in regard to regions, states or cities within a specific geographical area or a particular country.

1.6 Need for the research

Promoting the facility location decision-making process to help founders of new entrepreneurial firms to choose the best-fit location, along with developing a list of critical

factors that most likely attract entrepreneurs to potential locations has multiple advantages for both entrepreneurs and regional development authorities.

- Advantages to entrepreneurs
    - More reliable decisions,
    - Creation of greater wealth,
    - Achieving self satisfaction both personally and professionally, and
    - Better understanding of new and different cultures
- Advantages to regional development authorities
    - Economical development through adding to the national income of the country generated from establishing new businesses through:
        - Payment of business registration fees,
        - Expenditures on patent-related components,
        - Rental or purchasing business spaces,
        - Utilization of public services,
        - Generation of additional taxes, etc.
    - Social development, through:
        - Introducing novel goods and services that promote life style and ease of performing frequent tasks,
        - Contributing in the reduction of the unemployment rates via providing direct and indirect job opportunities,
        - Elevating the education level to cope with requirements of a new life style or needed qualifications,
        - Participating in charitable activities and society diversification.

CHAPTER TWO

Literature Review

2.1 Entrepreneurial facility location literature review

The main goal of entrepreneurs across various industries is to mobilize all possible means to insure the ultimate success for their fledging ventures. To do so, the entrepreneurs when forming new ventures, encounter crucial strategic choices about resources, products/markets and activities (Manolova, Brush & Edelman, 2011). One distinct choice that they are required to handle at the early stages of their activities is where to establish the new venture, i.e., the location decision of the entrepreneurial firm.

From a firm size perspective, large firms have the advantages of scale, experience, brand name recognition, and market power (Chen & Hambrick, 1995). The small firms, however, need to be located where a pool of resources, a higher range of opportunities, and a lower rate of threats can be secured. Furthermore, the entrepreneurs usually operate in an environment of substantial and social ties that affect the start-up process (Manolova, Brush & Edelman, 2011). Thus, choosing the best location is a critical decision that has great impacts on many future decisions because the optimal location reinforces the ability of the newly initiated venture to expand or grow and obtain a competitive advantage.

Another distinct difference between small and large firms in decision-making is their tendency to seek closer proximity to customers (Mazzarol & Choo, 2003). Because many small businesses have a relatively limited base, the industrial estates, to which small ventures are more attracted, arrange themselves in a pattern of having one or two large firms, around which a large

number of small firms then cluster, acting as suppliers to the larger firms. This process, in turn, secures constant demand for the small firms' products/services and expands their rate of success.

In general, the decision where to locate the entrepreneurial venture depends mainly on its owner(s)/manager(s) analysis, derived by personal motivation, the social environment and the external business culture (Nijkamp & Ommeren, 2004). In order to formulate better decisions, business owners seek updated information that is relevant to products/services introduced through the business. The needed information is mostly gathered by talking to customers, participating in conferences, and attending trade shows to keep up to date with customer needs, technological improvements, and to develop ideas to promote products and services (McCarthy, 2003). Also, the emerging information and telecommunication advancement has emphasized the spatial connectivity potential for many locations and provides more reliable data in favor of new and innovative activities.

Various studies have indicated that decision makers in firms consider, to a large extent, locations where the economic profit can be maximized (Espitia-Escuer, Garcia-Cebrian & Munoz-Porcar, 2014). Yet, empirical perception indicates that decision making agents when optimizing their location decisions do not choose a potential location based only on a single objective; rather, they consider a range of often conflicting objectives to determine a location fitting for the firm.

In a familiar environment (e.g., local or domestic regions), the entrepreneurs usually have fewer complications to overcome in identifying social and economic resources. This situation would strengthen their ability to establish more viable organizations. On the contrary,

unfamiliarity with the business environment in which to start the venture adds extra obstacles to secure the required resources and contacts.

On the other hand, choosing distant locations rather than founding the firm locally might enhance the accumulation of physical resources and mobilizing additional financial resources. Establishing the firm locally might be constrained by zoning ordinances, transportation access or physical size (Manolova, Brush & Edelman, 2011). Also, choosing a distant location for the firm gives it greater legitimacy, increases its acceptance as a separate entity and signifies the entrepreneur's tangible commitment to build the venture, which in turn, induces suppliers and outside financers to trust offering higher credit to distant firms than to their local counterparts.

Entering a foreign market is another critical strategic decision the organizations have to handle with great caution and elaborate investigation and research. Based on the economic and investment nature of the targeted market, firms (specifically small and medium enterprises) have to choose the most suitable entry mode to utilize for entering that market since the choice of a particular mode will be difficult to change and will cost valuable time and money.

There are four common entry modes to foreign markets exporting, licensing, joint venture, and sole venture (Agarwal & Ramaswami, 1992). According to normative decision theory, the entry mode into a foreign market is chosen based on trade-offs between risks and returns. Besides choosing the entry mode to foreign markets that has the highest risk adjusted return on investment, decision makers also look into resource availability through which the firm's financial and managerial capacity can be assessed for serving the targeted foreign markets. Decision makers in entrepreneurial firms take into account the need for control to influence systems, methods, and decisions in those foreign markets. Moreover, the determination of a

particular entry mode of foreign markets involves delicate adjustments of both firm and market factors that have major effects on the main four entry mode criteria risk, return, resources, and control (Agarwal & Ramaswami, 1992). If a firm chooses the exporting entry mode to decrease the associated degree of risk when entering a foreign market, most likely it will need to mobilize low investment (low financial resources). This strategy would also provide the firm with quite high operational control, but at the same time, its marketing control would be limited to generate influence in the targeted market. The licensing mode conveys the need to low investment and a low degree of risk, but it will only give the firm the least operational and marketing control. On the other hand, when the decision makers select the sole venture mode as their firm's entry strategy to a foreign market, the firm will be provided with a high degree of control, but this will be accompanied by the need for high investment and will include high risk and return. Finally, choosing the joint venture mode to enter the foreign market involves a relatively lower investment and provides a proportionate risk, return, and control.

Entrepreneurs are well known for their ambition, independence, self-confidence, and innovation. Among several other traits, they are also risk-bearing and strive for formal authority (James Carland, Hoy, Boulton, & Jo Ann Carland, 1984). To achieve their goals and satisfy their urges, the entrepreneurs usually align knowledge and resources to start small ventures. Thus, choosing the sole venture entry mode when starting their small businesses in any market is most appropriate to fulfill the desired criteria, including foreign markets.

Table (2.1) Summary of the literature review on entrepreneurial facility location

Author	Year	Concept	Contribution
James Carland, Hoy, Boulton, & Jo Ann Carland	1984	Choosing the entry mode to achieve entrepreneurial goals & satisfy entrepreneurship needs	Sole venture entry mode, to minimize financial risk and have greater level of control

Agarwal & Ramaswami	1992	Modes to enter foreign market	- Entry mode depends on trade-offs between risks and returns; - Enter markets that have available of resources, - Need for control to influence systems, methods and decisions - Influenced by adjustments of firm and market factors; risk, return, resources, and control
Chen & Hambrick	1995	Relation between firm size and choice of location	Small firms are preferred to be located where pool of resources, higher range of opportunities, and lower rate threats exist
Mazzarol & Choo	2003	Tendency of small firms to be located in proximity to customers	Small firms are located around one or two large firms
McCarthy	2003	Importance of obtaining adequate information for better location decision	Source to obtain information: talking directly to customers, participating in conferences, attending trade shows all supported by emerging information and technological advancement
Nijkamp & Ommeren	2004	Influence of personal motivation on location decision making	Small firms location decision making depends heavily on owner's analysis that is derived by their type of personality
Manolova, Brush & Edelman	2011	- Location decision is crucial for firms - It is more critical for entrepreneurial firms	- Choosing location is important since early stages of establishment - Making good location decision reinforces the expansion and growth to obtain competitive advantages
Manolova, Brush & Edelman	2011	- Advantages of locating the firm at distant locations - Limitations of choosing local sites	- Enhance accumulation of physical resources, mobilize more financial resources, gives greater legitimacy - Constrained by zoning ordinancess, transportation access, physical size
Espitia-Escuer, Garcia-Cebrian & Munoz-Porcar	2014	Factors to consider in location decisions for small firms	- A range of potential conflicting objectives - Maximizing economic profit

2.2 International entrepreneurship literature review

The decision of locating entrepreneurial firms in a foreign market (internationally) entails decision-making strategies and approaches that are anisotropic from those adopted for organizations that choose domestic or local regions as venues for their activities.

Due to the expected competition with local firms when the entrepreneurial firms choose to be located in a foreign market, these firms are required to mobilize sufficient assets, skills and resources to secure costs and fulfill demands associated with operating in the foreign market. Assets are needed to provide the firm with the necessary means to successfully compete with the domestic firms. For example, the lack of multinational experience, particularly the experience of the targeted foreign market, can lead to the exaggeration of involved risks. Specific skills are required to develop differentiated products or customized services to identify potential customers in the targeted foreign market, considering using a high control mode to prevent the loss of long-term revenues if knowledge/knowhow is shared with local firms. Well-integrated resources are also of high importance to obtain, if necessary, including related patents or collaboration contracts, and reducing marketing costs.

Moreover, developing sustainable competitive advantages is a fundamental part of the decision-making strategy for any firm to be able to create wealth, specifically those firms that have decided to go global or to be located in an international market. Several approaches help to formulate such strategies (Rialp-Criado, Galvan-Sanchez, & Suarez-Ortega, 2010) in which the level of control and integration; more predictable environments; implementation of the entrepreneur/founder's vision, experience, and knowledge, and a viable match between opportunities and threats exist in the external foreign market; the set of resources and capabilities of the organization; shared values and norms in the culture of the targeted market to provide a

guide to appropriate behavior; and responsiveness to different demands and conditions of the environment are embedded.

As entrepreneurship can be defined as the act of entry to markets, it is the entrepreneurial manager's responsibility to decide what markets to enter, the time of entry, and the entry mode and approach (Lumpkin & Dess, 1996). Similarly, the international entrepreneurship concept is implemented when the firm's business and activities cross national borders with the focus on a relationship between businesses and international environments they operate in (Wright & Ricks, 1994). International entrepreneurship is multi-disciplinary and is based on related theories from international business, entrepreneurship, economics, psychology, anthropology, finance, marketing and sociology (Oviatt & Mcdougall, 2005). To include undertaken risk as the defining act, the international entrepreneurship definition was further refined by (Mcdougall & Oviatt, 2000) as the combination of innovative, proactive, and risk-seeking behavior that crosses national borders with the intention to create value in firms. Moreover, since the entrepreneurial manager is the one who would be also making the decision, the international entrepreneurship definition could be broadened as innovative, proactive or risk-taking behavior of an actor to undertake cross-national border activity through the act of international market entry (Perks & Hughes, 2008).

There are two main labels that are often applied loosely to describe venture types in the international entrepreneurship (IE) realm (Cviello, Mcdougall, & Oviatt, 2011). Since the mid-nineties scholars have been using the two terms international new ventures (INV) and born global organizations (BG), interchangeably within the broader IE literature. In fact, the term INV was extracted in reflection to its counterpart's research in the international business (IB) field in which involved scholars often distinguish between international and global terms. The IB

researchers use the term 'international' for crossing borders of a single country while the term 'global' is used for being active in many countries or continents (Cviello, Mcdougall, & Oviatt, 2011). Accordingly, in IE literature the INV term is mainly defining ventures that have competed primarily in their own regional market or in a relatively limited number of countries. The BG term, on the other hand, is used when describing organizations with a genuine global focus. This distinction is reflected in the conceptual distinction between geographically focused start-ups and global start-ups. In contrast, the INV and BG have a distinctive commonality between the terms 'new' and 'born.' Therefore, new and young firms should be the focal of INVs and BGs studies and IE scholars should take in consideration that it is the firm's age that should be the major defining characteristic rather than its size or its scope of foreign operations. This is because size and scope of the firm are greatly influenced by how early and quickly it grows and internationalizes the activities from its foundation time (Cviello, Mcdougall, & Oviatt, 2011). Thus, it is important for researchers to clarify the life-cycle stage of the firm in the study of international entrepreneurship.

Traditionally, several studies suggest that firms usually become international after a long period of domestic establishment (Oviatt & Mcdougall, 1997, 1999). However, whereas many firms still internationalize in a slow, gradual, and evolutionary path, other newer and entrepreneurial ventures become global or international almost at the time of their establishment. This is most likely due to the rapid changes taking place in the global markets and industries, as well as the escalating orientations of entrepreneurs towards internationalization (Oviatt & McDougall, 1995, 1997; McDougall & Oviatt, 2000).

Table (2.2) Summary of the literature review on international entrepreneurship

Author	Year	Concept	Contribution
Wright & Ricks	1994	Concept of international entrepreneurship	Implemented when business and activities cross national borders with the focus on a relationship between businesses and international environments
Lumpkin & Dess	1996	Decision upon market to enter and time of entry and mode	It is the responsibility of entrepreneurial manager(s)
Oviatt & Mcdougall	1997, 1999	Timing to go international for firms	- Traditionally, after a long period of domestic establishment - More recent approach to go international almost at the time of their establishment
Oviatt & Mcdougall	1995, 1997, 2000	Timing to go international for firms	Decision to go international at earlier stage is derived by rapid changes in the global markets and industries and the escalating orientations of entrepreneurs towards internationalization
Oviatt & Mcdougall	2000	Concept of international entrepreneurship	Refined definition of international entrepreneurship: combination of innovative, proactive and risk-seeking behavior that crosses national borders with the intention to create value in firms
Oviatt & Mcdougall	2005	Concept of international entrepreneurship	A multi-disciplinary approach based on theories from international business, entrepreneurship, economics, psychology, anthropology, finance, marketing and sociology
Perks & Hughes	2008	Role of the manager as the location decision maker	Innovative, proactive or risk-taking behavior of an actor to undertake cross-national border activity through the act of international market entry
Rialp-Criado, Galvan-Sanchez, & Suarez-Ortega	2010	Importance of developing sustainable competitive advantages	Strategies of location decision for small firms with global orientation takes into account level of control, predictable environments, vision, experience, and knowledge implementation, and viable match between existing opportunities and threats

Cviello, Mcdougall, & Oviatt	2011	The two main labels for type of ventures in international entrepreneurship	- International new ventures (INV): defines ventures competing in their own regional market or in a relatively limited number of countries. - Born global organizations (BG): describes organizations with a genuine global focus

2.3 Strategies and factors to choose an international market literature review

Many publications on factors that are used as a basis for location decisions of enterprises in general fall into two broad categories (1) studies to measure the influence of a specific factor or a set of factors on firm location decisions, such as analyzing the impact of taxes and incentives, and (2) studies that explain the decision process for a specific business or industry, e.g., the location decision process of biotechnology firms (Kimelberg & Williams, 2013).

Scholars of location decision have continuously turned their attention towards the factors that influence the location decision patterns over the years based on the core activity of firms. In the early and mid-twentieth century, where manufacturing was the core activity of most businesses and firms relied on production and sale of goods to succeed and generate profits, more consideration was given to factors such as access to raw materials, transportation costs, labor costs, and access to markets. Later on and as costs remained a central concern in selecting the firm's location, more research has also explored the importance of other several factors, including taxes, financial incentives, unions and labor laws, and infrastructure. The shift to a postindustrial era and the emergence of a knowledge-based economy steered the attention of scholars towards a different set of factors such as the need of firms to get situated within networks of competitors and collaborators to capitalize on innovation and satisfying the

preferences and needs of current and targeted skilled human capital (Kimelberg & Williams, 2013).

Furthermore, the research on location selection adopts two basic methodological approaches (1) surveys of companies, and (2) statistical models. Surveys typically identify one or more key respondents and ask them about factors that influenced their location decision. Their advantages include reporting the stated significance of variables that are difficult or impossible to quantify and offering the ability to ask open-ended questions leading to perhaps the identification of unintentionally neglected factors. On the other hand, statistical models collect information and variables on new economic activity, such as the establishment of new plant and explore some of the factors that influenced the selection of a specific location. Such statistical models have the advantage of determining the size and direction of relationships among factors that would be difficult to obtain using the surveys (Carlson, 2000).

The increasing interest of small firms from the stage of their outset in internationalization and going global is derived from several internal and external key factors and trends (Rialp-Criado, Galvan-Sanchez, & Suarez-Ortega, 2010). New development of market conditions in many sectors of economic activities, technological revolutions in production, transportation, communication, etc., global networks and alliances' prosperity, and the growing number of skilled people with entrepreneurial orientation (Rialp et al., 2005a, 2005b) are among most common factors that encourage the phenomenon of born global firms.

Changes in market conditions are rapidly encouraging the establishment of small ventures with flexible and dynamic internationally oriented business operations. In spite of their limited resources, small firms adopt more specialized production and operations strategies to serve specific niches in the international markets that have deficiencies in meeting their

customers' demand. They also depend on their distinctive competencies to produce innovative and distinguished products that can be sold worldwide (McAuley, 1999) and therefore reinforce their capability to compete with local competitors.

Recent technological improvements help small firms to generate profits in the international markets through several aspects. Issues such as specialized production and client adoption are more viable for small-scale operations due to improvements in manufacturing technologies. Advanced transportation offers more reliable, frequent, and cheaper means of movements between countries and continents and therefore cuts the cost required for moving people and goods.

Development of information technology has allowed easier data accessibility and collection as well as simplified the data analysis and interpretation. This technology has provided entrepreneurs with more tools to identify new opportunities and circumstances that in return enable them to carry on planning managing international activities from the time of their venture's founding (Oviatt & Mcdougall, 1994, 1995, 1997).

A growing number of manufacturing and services sectors are improving their cross borders networks and links through creative procedures of global supply and distribution (Jones, 1999) in more rapid integrated global markets.

In the last decades, people with entrepreneurial orientation tended to acquire more skills and obtain more international education and experiences (Andersson, 2000), which has increased the number of small firms' founders who can interact and negotiate with entities from different cultures and therefore take their small ventures internationally more frequently.

Other important external factors of foreign business environments that might be of high attractiveness to entrepreneurs with internationalization orientation and that contributed in explaining the rise of international new ventures are the internalization of transactions, an alternative governance structure, the development of the foreign location advantage, and a unique resource control (Cannone, Costantino, Pisoni, & Onetti, 2012).

Such similar factors stimulate the rapid engagement of small firms with multiple international markets from the inception stage via utilizing global networks that help to align extra resources for cross borders' outreach (Rialp et al., 2005b).

There are specific internal factors that are of great influence on international market entry. These factors are related to entrepreneurs or managers since they are the primary (in many cases the sole) decision makers of newly established small firm. Actually, individual characteristics and attributes of the entrepreneur, such as previous international and business experience, academic training, ambition and motivation levels, risk perception, global vision, leadership ((Oviatt & McDougall, 1995, 1997; Madsen & Servais, 1997; Rialp et al., 2005b), and personal relationships (Madsen & Servais, 1997; Harris & Wheeler, 2005; Rialp et al., 2005a; Gabrielsson et al., 2008) are critical variables to formulate the necessary strategies for the firm to go global. Therefore, distinctive and sophisticated entrepreneurial capabilities of the small firm's founder play a key role to make advantage of the considered international market opportunities.

To obtain higher returns when operating and servicing foreign markets the interested founders of small firms are expected to use a selection strategy and favor entry into more attractive markets. The market attractiveness is most likely characterized in terms of the market potential (size and growth) and the associated investment risk, which have been found to be

important determinants of overseas investment (Forsyth, 1972; Weinstein, 1977; Khoury 1979; Choi, Tschoegl and Yu, 1986; Terpstra and Yu, 1988).

The international entrepreneurial culture (IEC) within firms with global orientation also has received considerable interest from IE scholars as another significant internal factor. It concerns the international entrepreneurial activities of the firm to identify and pursue opportunities abroad (Dimitratos & Jones, 2005; Zahra, Korri, & Yu, 2005). Typically, IEC that assists globally oriented firms in their thriving to explore opportunities in the international markets consists of six interrelated organizational culture dimensions international market orientation, international learning orientation, international innovation propensity, international risk attitude, international networking orientation, and international motivation (Dimitratos & Plakoyiannaki, 2003; Zahra et al., 2005).

International market orientation consists of international customer orientation, inter-functional coordination, and international competitor orientation (Narver & Slater, 1990). It describes the firm's adopted behavior to provide its foreign customers with superior value in products or services. This orientation requires changes in the organizational culture of the firm to cope with the requirements of operating successfully in the international markets. Accordingly, the existence of a strong international market orientation within the entrepreneurial firm facilitates going global and entering international markets (Armario et al., 2008; Perks & Hughes, 2008); increases knowledge-creating capability abroad (Nguyen & Barrett, 2006); and enhances the international performance (He & Wei, 2011; Knight & Kim, 2009; Kropp, Lindsay, & Shoham, 2006; Racela, Chaikittisilpa, & Thoumrungroje, 2007).

International learning orientation is a significant characteristic that is embedded in the organizational culture with international oriented firms. It helps the entrepreneurial firm to

explore international market trends and demands that could be carried out through three processes (Moorman, 1995; Nevis, DiBella, & Gould, 1995) information acquisition, information dissemination, and information use. International learning orientation can also affect greatly the firm's capability to identify business opportunities in the international market and therefore define its business scope (Voudouris, Dimitratos & Salavou, 2012).

International innovation propensity is an integral part of the organizational structure of the entrepreneurial firm (Lemon & Sahota, 2004) to support new and creative ideas, products, and processes that are specifically laid out for foreign markets (Knight & Kim, 2009). The organizational decision upon innovation has a direct impact on the firm's internalization intention and thus it is important for the same to be taken at the stage of its establishment.

International risk attitude is another major component of the organizational culture that is related to the readiness degree of the firm to get engaged in substantial and risky resource commitments in international markets (Miller & Friesen, 1978). It allows the internationalized firm to favor low to high-risk alternatives, gradual to wide-ranging behavior, and conservative against bold decisions in situations of uncertainty (Khandwalla, 1997).

International networking orientation, which is a part of the organizational structure of the internationalized entrepreneurial venture, promotes its capability to actively operate in cross-borders through identifying resources in the external environment and forming alliances (Granovetter, 1985; Gulati, 1998). International networking is associated with both international market orientation in the sense of developing the business-to-business relationship (Gellynck, Vermeire, Viaene, 2007) and learning that could be occurring from business and social networks formed with domestic agents (Freeman, Hutchings, Lazaris, & Zyngler, 2010). Innovation efforts

of the internationalized small firm may also be motivated by an efficient international networking (Kelley, Peters & O'Connor, 2009).

International motivation requires distinct administrative approaches (Zhou et al., 2006) and building internationally oriented management teams within the organizational structure of the entrepreneurial firm. It is important to assist the decision maker to clarify available business opportunities abroad (Zahra et al., 2005) for born global organizations and multinational enterprises (Lee & Williams, 2007). Moreover, international motivation urges a learning process (Michailova & Minbaeva, 2011), as well as the formation of international networking (Lee & Williams, 2007) in the multinational and global firms.

Knowledge, as a stand-alone factor, is considered the most influencing internal factor. Decision makers depend heavily on the accumulated amount and type of knowledge the entrepreneurial firm possesses to determine which model of internationalization to adopt the incremental internationalization model, the born global firms model, or the non-sequential internalization model. The types of knowledge relevant for the internationalization process include knowledge about how to manage increased complexity and diversity in international markets, knowledge of the foreign markets, clients, and competitors, and knowledge of foreign government institutional frameworks, rules, norms, and values (Cuervo-Cazurra, 2011). Additionally, variation between companies in using their knowledge base results in the existence of various internalization processes.

(Beckers & Kloosterman, 2011) in their UNU-MERIT working paper contrasted two pre-WWII and post-WWII business neighborhoods within Dutch regions. After reviewing the zoning regulations through group and individual interviews of these neighborhoods' experts and

entrepreneurs, they argued that founders of migrant businesses are motivated to locate their ventures in specific neighborhoods based on more factors than sheer costs and benefits. Factors such as knowledge and available information about rules and regulations of the region that are related to a particular line of business have significant importance to obtain the necessary start-up and social capitals, including providing linkages to local suppliers, customers and labors. Furthermore, their thorough research of how the size and the cost of business spaces are influencing the location decision of migrant entrepreneurs resulted in identifying key dimensions that offer business opportunities and demographic characteristics that partly shape potential supply as well as the demand of products (Rekers and van Kempen, 2000), the built environment with its local policies and supporting regimes (Ram et al., 2002), the increased tendency towards self-employment (Schutjens and Stam, 2003; Stam, 2009), the increased outsourcing of business activities by large firms, the rise of internet commerce, and the growing flexibility of labor contracts (Wennekers et al., 2008).

Various literature embedded in the (PBL) study that has been conducted by the Netherlands Environmental Assessment Agency during the year 2010 has identified location decision factors that affect business functioning: accessibility and parking (Hegens et al., 2009), local market prosperity (Bulterman et al., 2007; Ouwehand and Van Meijeren, 2006), local livability and business location image (e.g., the crime rate, the status of the built environment, vandalism, and dirty public spaces) (Wilson, 1987), the presence of certain local amenities (McCann and Folta, 2008; Florida, 2002; Weterings et al., 2009), and the availability of local business spaces (Aalders et al., 2008). Through interviewing forty local entrepreneurs, the (Beckers & Kloosterman, 2011) paper has also determined five blocks of key location factors that affect business operations:

- Cost-saving potential, including reasonable business spaces,
- Market potential, including cleanliness, safety, firm accessibility, and parking space availability,
- Local social embeddedness,
- Convenience, including the firm's proximity to the entrepreneur's home,
- Firm spatial needs,
- Contacts with local residents, and
- Ethnic population mix.

(Sinkiene and Kromalcas, 2010) wrote an article on the concept, directions and practice of city attractiveness improvement as a part of a public policy and administration report in Lithuania. In the article, they stated that there is a shift in efforts concerning city (location) development from heavy industry to creative, talented and highly skilled activities. For the location (city) to be in a better competitive position, various internal and external factors must be the focal point, such as a highly skilled labor force, creative entrepreneurs and workers, clean and high value-added businesses that are the engines of knowledge economy and therefore stimulate the locational economy. Moreover, the international competitiveness of a territory increases due to critical determinants, including processes of democratization, decentralization, transfer of decision-making power, development of information technologies, and free movement of people, capital and goods, for which governments have to initiate and implement complex strategies.

Other factors attracting businesses' representatives to a location include the labor market quality and size, as well as the quality of the residential environment (Berg, Meer, and Otgar, 1999). The rapid globalization phenomenon is urging business environments to explore means to

increase their competitiveness competencies at national and international levels, which in turn lead to a new approach of cities-entrepreneurs that use available resources to acquire higher competitiveness in the economic, social, or environmental fields (Kotler, 1993; Porter, 1998), allowing them to choose the optimum location without any barriers. According to Gorzelak, 2001, choosing the best location for the entrepreneurial firm has been influenced by the shift in the twenty-first century towards knowledge-based economies where the markets demand a creative and complex workforce. Critical factors of business attractiveness are categorized into two distinctive groups:

- Factors related to a resource-based economy labor force, resources, premises, bulk transportation, and energy resources.
- Factors related to a knowledge-based economy qualification, research and development centers, local supplies, reliable infrastructure, and good living conditions.

In particular, attractiveness of the market in the knowledge-based economy to entrepreneurs is determined by distinctive factors: a friendly and stable attitude, effective and honest promotion, competition in capital, innovations, and labor.

Furthermore, (Berg, Meer, and Otgar, 1999) defined the factors that boost the city's (location's) attractiveness to include good accessibility, reasonable land prices, local taxes and legal requirements, sufficient quantity and quality of the labor force supply, market size, city (location) status, living environment, and the quality of public services. From his perspective, (Braun, 2008) suggested that there are important characteristics of the city (location) that entrepreneurs and investors look for, including: location, built environment, labor force, existing and new customers, suppliers, and financial partners.

Local governments in many countries consider creating protected areas that are provided with adequate infrastructure and easy access to local resources along with restricted policies to protect them from misuse and harmful exploitation to stimulate the development of the regions around the country (Hammer, 2007). However, the challenge is emphasized in identifying and promoting business opportunities that do not prevent the utilization of the protected site's resources while, at the same time, taking into account minimizing the resulted negative impacts. Therefore, there is no doubt that resources and values of the natural environment affect the business environment as an attractive location for new companies. In such locations covered by legal protection, attracting factors such as information and promotional support, grants and subsidies, the advice of business environment institutions, the use of exemption and tax benefits, and assistance in financing as well as in adjusting the profile of requirements to operate in the location, have been revealed by entrepreneurs to be of very high importance (Analiza, 2012).

Empirical studies that have been conducted in 2011 in 229 rural communes in the Mazowieckie region in Poland revealed that local authorities implemented strategies aiming mainly at attracting outside businesses along with supporting local entrepreneurship through considering improvements of the social and technical infrastructure. These studies have also provided evidence on how entrepreneurship development is an essential element to promote the local economy, which is reflected as an increment in GDP per capita, job growth, and positive changes in the economical structure of the studied areas (Golasa, 2015). There were also other determinants that are associated with the areas' attractiveness to new businesses including an increased number and quality of services and resources (e.g., developed land, real estate, etc.), as well as intellectual resources (skills, knowledge, and qualifications of local community members) (Struzycki, 2006). Moreover, enhancing regional attractiveness for new businesses

requires local authorities to introduce ambitious plans to provide the basis for creating optimum features for investors, optimizing the use of the limited financial resources and assisting businesses to secure financing from external sources, better adaptation to environmental changes (arising opportunities or threats), and conducting promotional activities.

In their study (Hui Tseng, Tansuhaj, Hallagan, & McCullough, 2007) on antecedents of multinational expansion, the authors have noted that foreign expansion demands assessing more selective resources to help buffering the associated costs and risks with moving to international markets. This is mainly because taking the business abroad involves greater managerial complexity and liability of foreignness. The study has also found out that research and development intensity is highly important for firms' expansion behavior across borders (Davidson and McFetridge, 1985; Gatignon and Anderson, 1988; Chen and Hennart, 2002). Additionally, four categories of the firms' resources developed locally would be highly beneficial when several location-specific advantages are offered by host countries, including labor availability, production facilities, and distribution channels (Chen, 2005). One category is the technological competencies, because they have a collective good characteristic to be replicated without incurring full costs (Caves, 1971, 1996; Martin and Salomon, 2003). A second category of resources firms seek to benefit from globally is their marketing resource. Marketing advantages of strength of brand image, achievement of scale economies in marketing, and owning bargaining power with distributors and consumers could overcome the cross-cultural differences and help to mobilize consumer preferences and enhance marketing environments and infrastructures. The third category of resources is property-based. The organizational flexibility helps the firm to cope with the more global integration of the business community that can be seen in the mobility of some parts of the value chain into different places where they can perform

more efficiently or create greater value. It is the firm's flexibility of resources that allows its corresponding move with its suppliers or customers to be feasible and therefore to compete in international markets with less binding constraints. Another category of resources is the financial resources that influence the firms' behavior internationally. The ability to raise financial support externally through capital markets or financial institutions in the foreign market is a very important factor of attraction to businesses.

The shift in more emerging economies towards being knowledge-based (e.g., South East Asia and Eastern Europe) has been supervened by a significant decrease in the brain drain phenomenon and has resulted in threatening countries (USA, for example), that depend heavily on foreign talent and competiveness as a land of opportunity (Mahroum, 2000). Actually, governments can play a chief role in providing incentives for foreign talent to stay in the country or to move abroad. Countries that desire to be in top shape for talent must have attractive governmental policies for the intellectuals and innovators such as providing tax incentives, superior research infrastructure, and competitive compensation structures to attain their skills and encourage more talent to come. Consequently, their business environment will become more attractive for new firms. The nature and structure of a national innovation system (NIS) of the country can also impact the inflow of highly skilled people; countries with NIS that is based on its universities' capabilities will most likely attract academics, whereas other countries of high foreign direct investment (FDI) provide more incentives to expatriate professionals who move along within their companies. Other governmental regulations of some countries, including visas, taxation, and protection, along with credits for facilities, stimulate entrepreneurs to immigrate and settle in these countries. On the other hand, governments have to consider factors that might be detrimental to entrepreneurial activities such as bureaucracy, an unfavorable entrepreneurial

climate, inflexible human resource management (hiring, firing, work hours), and a lack of available venture capital (Mahroum, 2000).

In his paper (Garza, 2012) argued that there is a major trend among the international companies to move their activities to developing countries for cost savings (Huang, Zhang, Zhao, & Varun, 2008). These developing countries offer cost savings in the form of lower factory wages along with other attractive business environmental factors such as favorable exchange rates, a significant amount of unskilled labor, and favorable foreign trade policies. In fact, the cost behaviors of the firms must be analyzed extensively to take corrective actions if necessary in the contexts of the short life cycle of products and the rapid increase in global competitive pressure. The manufacturing costs could be lowered by minimizing the involved costs of some or, if possible, all of the components that comprise the total cost of the product. As per the Kearney attractiveness index developed on 2004, there are three primary drivers for offshoring (taking the business across borders):

- Financial factors that include
    - Compensation costs (average wages and median compensation costs),
    - Infrastructure costs (costs of occupancy, electricity, and travel,
    - Tax and regulatory costs (relative tax burden, costs of corruption, and fluctuating exchange rates.
- Workforce skills and availability that include
    - Cumulative business process experience and skills (existing market size, and quality rankings of management training,
    - Labor force availability (total workforce, and total educated workforce),

    - Macroeconomic variables (relative country economic growth, and unemployment rates), and
- Overall business environmental factors. These include
    - Country environment (overall business and political environment, and the extent of bureaucracy),
    - Country infrastructure (blended metric of infrastructure quality),
    - Cultural adaptability (personal interaction score (extracted from Kearney's globalization index)), and
    - Security of intellectual property (investor ratings of the IP protection, and software piracy rates).

In spite of these new businesses-attraction determinants, there are six major counter factors (Garza, 2012) that lead to reshoring (bringing the business activities back home): increasing wages of local workers in the developing countries, and the rise in shipping costs to reach the final customer, the rapid elevation of inventory costs (affected by long lead times that force the companies to keep a stock of at least three extra months of inventory (Koepfer, 2011)), quality control cost (for which some companies have to hire personnel who are totally devoted to control the quality of incoming shipment from the offshored country, adding more cost on each unit, besides the growing awareness of customers about the 'made in' label and additional time and cost to send the products back), prototyping expenditure (prototyping is carried out through the research and development department at the country of origin, increasing the production costs, whereas working closely with the production managers and assembly workers trims the costs per unit) (Davidson, 2010)), and the intellectual property protection costs (the less

vigorous enforcement of intellectual property protection laws costs companies hundreds of thousands in lawsuits to protect their patents (Sherwood, 2000)).

One prominent feature of the successful international market concept is providing new businesses with particularly attractive locational factors that ensure the utilization of most embedded resources within these incoming firms (Dunning, 2009). Companies with knowledge-intensive assets usually seek hosting locations that contain an abundance of skilled labor and a good public infrastructure. Also, new companies always favor business environments with fewer natural and artificial trade barriers and transaction costs. Another appealing locational factor is the ease by which firms are able to coordinate their cross-border activities and mobilize alliances with other local and foreign firms. Furthermore, economic and institutional facilities of the location, such as the existence of other foreign investors and the presence of a business cluster that offers specialized support services, are increasingly valued much higher than traditional criteria by multinational enterprises (MNEs). Promotional campaigns and incentives in the form of a short process of planning applications, land grants, subsidized rents, tax holidays, and generous investment allowances, as well as the macro-economic or country-specific characteristics of the distribution of natural resources, specialized labor, and the availability of land and finance capital, introduce further attractive variables of the location.

A study conducted by Gorter in the year 2000 on migrant entrepreneurs in East Indonesia indicated that both economic and non-economic factors can determine the location attractiveness to entrepreneurs. Institutional support is highly important for the innovation processes of entrepreneurs (Schumpeter, 1934). To pull entrepreneurs towards a business location the credit institutions present in that location (e.g., banks) should have efficient procedures to provide funds for entrepreneurial firms to carry out their activities. Socio-economic conditions of the

location that are important include the level of competition, access to markets, access to capital, and the availability of information about the local tolerance degree, the existing supporting networks, and the niche concentration (Mulligan & Reeves, 1983; Gouch, 1984; Timmermans, 1986). The study also discussed non-economic and social factors of the location attractiveness; firms must be successful in coping with the culture of people in the foreign market in order for them to survive in that location (Waldorf, 1994). New venture founders also have to consider the religion in some countries when assessing the potential location. Other factors such as economic crises and insecurity of business activities also are substantial determinants of locational attractiveness.

In their location decision process, the entrepreneurs tend to choose locations based on the principle of profit maximization and risk minimization. From their side, municipalities have to reinforce the strategies that help to simplify the process of starting and running business activities (Jarczewski, 2008). Some of the basic actions municipalities could adopt to attract new businesses are preparation of real estate with the provision of the physical plan, technical infrastructure and accessible roads, real-estate tax exemptions, and attracting large investors that would most likely promote the goodwill and pro-investment image of the location and consequently accelerate the influx of other businesses.

International entrepreneurship (IE) literature in its two main labels of international new ventures (INV) and born global (BG) addresses many issues that are related to the entrepreneurial firms' endeavor to internationalize their operations and activities. The IE scholars have identified four main categories in their arguments around small entrepreneurial enterprises with international orientation (1) individual entrepreneurs, (2) the entrepreneurial process, (3) environmental factors, and (4) smaller entrepreneurial ventures, with only a few or infrequent

studies exploring the importance of the environmental factors in the entrepreneurial organizations' location decision-making process (Szyliowicz & Galvin, 2010). In general, IE research provides the entrepreneurs, founder, or decision makers with a relatively accurate supplementary tool to choose the best international market to enter even before the stage of the entrepreneurial firms' establishment.

Typically, firms deal with their location decision in a two-stage process. In the first stage the site requirements and the relative importance of these requirements are established. Then, in the second stage, the determined criteria are applied to the candidate sites in order to eliminate unqualified locations until the most suitable sites are identified. In fact, most of the firms consider the location decision as a final step in the macro-economic analysis part of their feasibility study, in which a preliminary screening is executed to nominate potential geographical areas, followed by evaluation of some of these areas to narrow the number of alternatives to choose from based on related location factors (Yang & Lee, 1997).

The comparative evaluation of potential locations through examining related location factors can by carried out through both traditional mathematical models such as mixed integer programing and decision analysis as well as various new facility location decision models including simulation models, expert systems, and neural network techniques. Furthermore, several of these approaches can be used for multi-criteria decisions, such as the analytic hierarchy process (AHP) and the multiple criteria decision analysis (MCDA) methods, specifically, the value measurements models (e.g., Multi-attribute Value Theory (MAVT)).

The analytic hierarchy process (AHP) is one of the most frequently applied methods for decision support. The process is based on a hierarchical decomposition of decision problems into

multiple criteria and the preferences are assessed using pairwise comparisons. The aggregation of these pairwise comparisons is then applied into the overall evaluation of considered alternatives within the decision problem (Durbach, Lahdelma & Salminen, 2014). The analytic hierarchy process (AHP) and the analytic network process (ANP), which is a general form of (AHP), involve four steps: (1) decomposing the problem into set of hierarchical or network models; (2) generating pairwise comparisons to assess the importance of considered elements; (3) developing a matrix by which the priority of elements is represented; and (4) taking decisions based on the pairwise comparisons matrix (Yang, Chuang, Huang and Tai, 2008).

The location selection model using AHP/ANP is a three-step procedure: determining initial criteria; identifying detailed criteria; and implementing an evaluation model. The procedure and its steps are further illustrated by an example of choosing a profitable location for a shop operating in the service industry.

Step 1. Building the initial model: the initial model for selecting a profitable location for a shop in a service industry is to be built with the consideration of three main criteria: market attraction; consumer characteristics; and location qualifications.

Step 2. Identifying the detailed criteria: appropriate dimensions and detailed criteria for choosing potential locations for the shop are identified through the judgment of six external experts in shop location selection and chiefs of marketing and sales departments. Any dimension that scores an importance of more than 90% in the reviewers' judgment will be listed among the detailed criteria, while criteria scoring an importance range between more than 70% and less than or equal to 90% shall be discussed further with reviewers to give final decision upon listing them among the detailed criteria.

The final list that will be considered in the evaluation model contains three main dimensions with their eighteen detailed criteria:

1. Market attraction: passerby flow, security issue, clustered market, public transit, and competition.
2. Consumer characteristics: consumer populations, consumer density, disposable income, purchasing power, and brand loyalty.
3. Location qualifications: rent, flexibility of lease term, shop size, employee recruiting, expected revenue, visibility of the shop, accessibility of the shop, and synergy between each branch.

Step 3. Implementing the evaluation model: the importance of the relationship between the dimensions and detailed criteria is judged by performing pairwise comparisons. The pairwise comparisons are conducted through separate questionnaires that are prepared in order for four marketing managers and two sales managers to utilize their experience in weighting the dimensions and criteria into two levels. Level one considered the comparison of criteria to determine which to be emphasized in the location selection for which a scale ranging between 1-9 is applied (e.g., Table 2.3a). Level two was used to compare the contribution of the dimensions by a scale ranging between 1-6 (e.g., Table 2.3b).

Table (2.3a) Representation for the criterion pairwise comparison in AHP

	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	
Market attraction						X												Consumer characteristics
Market attraction										X								Location qualifications
Consumer characteristics															X			Location qualifications

Table (2.3b) Representation for the dimensions pairwise comparison in AHP

	Market attraction	Consumer characteristics	Location qualifications
Market attraction	1	4	1/3
Consumer characteristics	1/4	1	1/6
Location qualifications	3	6	1

After comparing each of the elements, a paired comparison matrix (A) is formed and can be defined by

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix},$$

where a is the scalar value given to each criterion to be compared and n is the matrix order.

Since the pairwise comparison relies heavily on human judgment, there is a need to examine the consistency property of the pairwise comparison through the following:

1. Identifying the normalized pairwise comparison matrix A1

$$A_1 = \begin{bmatrix} a'_{11} & a'_{12} & \dots & a'_{1n} \\ a'_{21} & a'_{22} & \dots & a'_{2n} \\ \dots & \dots & \dots & \dots \\ a'_{n1} & a'_{n2} & \dots & a'_{nn} \end{bmatrix}, \; a'_{ij} = \frac{aij}{\sum_{i=1}^{n} aij} \text{ for } i, j = 1, 2, \dots, n.$$

2. Calculating the eigenvalue and the eigenvector

$$W = \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_n \end{bmatrix}, \; w_i = \frac{\sum_{i=1}^{n} aij}{n} \text{ for } i = 1, 2, \dots, n, \text{ and } W' = AW = \begin{bmatrix} w'_1 \\ w'_2 \\ \dots \\ w'_n \end{bmatrix}, \text{ and}$$

$\lambda max = \frac{1}{n}\left(\frac{w'1}{w1} + \frac{w'2}{w2} + \frac{w'3}{w3}\right)$, where W is the eigenvector, w_i is the eigenvalue of criterion i and λ_{max} is the largest eigenvalue of the pairwise comparison matrix.

3. Checking the consistency property

The consistency ratio CR = $\frac{CI}{random\ index}$, and CI = $\frac{\lambda \max - n}{n-1}$.

A set of recommended random indexes (RI) is given in the table below

Table (2.3c) Recommended random index (*RI*) by Saaty

N	2	3	4	5	6	7	8	9	10
RI	0	0.52	0.89	1.11	1.25	1.35	1.40	1.45	1.49

Table (2.3d) The eigenvector and the consistency ratio (CR) value

W	*W'*	
0.274	0.835	$\lambda_{max} = 3.054$
0.087	0.262	CI = $\frac{3.054-3}{3-1} = 0.027$
0.639	1.982	CR= $\frac{0.027}{0.52} = 0.052$

If CR is less than 0.1, then the comparison matrix is consistent. The eigenvectors and consistent ratios of the comparison matrices for detailed criteria in accordance with their upper level dimensions are given in Table (2.3e).

Moreover, the interdependence characteristics among elements and components can be handled through a supermatrix. The relative importance weight of each criterion from pairwise comparison is entered into the unweighted supermatrix (Table (2.3f)). Due to that the columns of the unweighted supermatrix sum to 1 and the components in the weighted supermatrix do not need to be weighted to make its column sum to 1.

A limit supermatrix, which is obtained by raising the weighted supermatrix to powers by multiplying it by itself, is used to determine the final local priorities to the global priorities. The limit matrix is reached and the multiplication process is halted when the column of numbers is the same for each column.

The limit matrix, shown in Table (2.3g), indicates that the most important dimension is the market attraction with a percentage weight score of 54.1% and the next most important dimension is location qualifications with a weight score of 34%.

Table (2.3e) The relative importance weights of the detailed criteria

Market attraction	Passerby flow	Security issue	Clustered market	Public transit	Competition
Weights	0.286	0.069	0.386	0.165	0.094
CR = 0.099					
Consumer characteristics	Consumer population	Consumer density	Disposable income	Purchasing power	Brand loyalty
Weights	0.070	0.114	0.294	0.384	0.138
CR = 0.093					
Location qualifications	Rent	Flexibility of lease term	Shop size	Employee recruiting	
Weights	0.080	0.044	0.053	0.029	
Location qualifications	Expected revenue	Visibility of the shop	Accessibility of the shop	Synergy between each branch	
Weights	0.343	0.133	0.205	0.113	
CR = 0.097					

Table (2.3f) Unweighted matrix and CR values

	Market attraction	Consumer characteristics	Location qualifications	Goal
Market attraction	0	0.750	0.833	0.270
Consumer characteristics	0.167	0	0.167	0.085
Location qualifications	0.833	0.250	0	0.645
Goal	0	0	0	0
CR Values	0	0	0	0

Table (2.3g) The limit supermatrix for dimensions

	Market attraction	Consumer characteristics	Location qualifications	Goal
Market attraction	0.541	0.541	0.541	0.541
Consumer characteristics	0.119	0.119	0.119	0.119

Location qualifications	0.340	0.340	0.340	0.340
Goal	0	0	0	0
CR Values	0	0	0	0

Considering three hypothetical locations A, B, and C where the data is collected from ten expert marketing and sales managers by direct questioning and questionnaires, the measures in the evaluation model are determined by applying a five-point scale that has integer values between 1 and 5; 1 is (low), 3 is (moderate), and 5 is (high) and the even values are for between the levels.

The weights of each detailed criterion (DCW) are obtained by using the AHP approach (Table (2.3e)), whereas the weights of each dimension (CW) are determined by the ANP approach (Table (2.3g)). The following table shows the mean score at each location.

Table (2.3h) Mean scores of each shop location

Criteria	Weights (CW)	Detailed criteria	Weights (DCW)	Scores		
				A	B	C
Market attraction	0.541	Passerby flow	0.286	3.0	3.0	3.0
		Security issue	0.069	4.6	2.0	3.8
		Clustered market	0.386	3.0	4.0	3.8
		Public transit	0.165	3.0	3.6	2.6
		Competition	0.094	3.4	2.4	2.4
Consumer characteristics	0.119	Consumer populations	0.070	3.0	4.0	2.4
		Consumer density	0.114	3.0	3.2	3.4
		Disposable income	0.294	2.8	2.6	3.0
		Purchasing power	0.384	3.8	3.4	4.0
		Brand loyalty	0.138	4.0	2.8	3.2
Location qualifications	0.340	Rent	0.080	4.6	3.0	3.8
		Flexibility of lease term	0.044	2.8	3.4	3.2
		Shop size	0.053	2.6	2.8	4.1
		Employee recruiting	0.029	2.4	3.6	2.6
		Expected revenue	0.343	3.0	3.0	3.0
		Visibility of the shop	0.133	2.4	3.1	3.4
		Accessibility of the shop	0.205	3.0	3.0	3.4
		Synergy between each branch	0.113	2.8	4.0	2.0

The overall result of each of the selected location is calculated as follows:

1. The detailed criterion score (DCS) is combined with a total weighted score of each of the main dimension (TSD) using the formula

   $TSD_{ij} = \sum_{j=1}^{6} \sum_{k=1}^{m} DCS_{ijk}\, DCW_{jk}$, where

   TSD_{ij} is the total weighted score of the dimension j and j = 1, 2, …, m of the evaluated location i.

   DCS_{ijk} is the score of detailed criterion k of dimension j of the evaluated location i.

   DCW_{jk} = the weighted value of detailed criterion k of the dimension j.

   i is the number of the evaluated locations (i = 1, 2, 3).

   j is the number of dimensions (j = 1, 2, 3, 4, 5, 6).

   k is the number of the detailed criteria (k = 1, 2, …, m).

   m is the total number of a detailed criterion k with respect to one of the upper dimension j.

2. The final weighted score for overall result (OR) is formulated as

   $OR_i = \sum_{j=1}^{6} TSDij\ CWj$, where

   OR_i is the weighted score of the overall result of the evaluated location i.

   CWj is the weighted value of the criterion j.

   The location that shall be selected is the one of the highest scores (Yang, Chuang, Huang and Tai, 2008).

Other multiple criteria decision analysis (MCDA) models also introduce better and less controversial decisions. MCDA aids the decision makers to organize and synthesize complex and conflicting information by taking explicit account of intangible criteria. Through MCDA,

objective measurements and value judgment are integrated together resulting in better exploitation and management of inevitable subjectivity (Beim & Levesque, 2003).

These and similar approaches can be successfully used in a wide range of applications including marketing, finance, education, public policy, economics, medicine and sports. Moreover, one of the main reasons for which they have been developed is to provide application-oriented solution procedures that can handle the involved complexity in large real-world problems. They also present a suitable substitution to the majority of available location decision models that do not take into account the qualitative location decision factors and that are deterministic in nature making them incapable to deal with rapid changes in the decision problems (Yang & Lee, 1997).

Although such methodologies can introduce a better solution to the large and complex location decision problems, in fact they still lack in overcoming considerable drawbacks such as quantifying all related location decision factors, decreasing the reliance on surveys and questionnaires in which the human judgment and opinion play a major role in the application of these methodologies, or the ability to comprehensively compare large number of possible sites. Furthermore, the rapid changes in the economy of the world along with the extraordinary revolutions in communication and debriefing means demand more convenient flexibility in adding or removing the decision making factors that are considered when choosing the best-fit location, which is not yet available in the current literature.

Table (2.4) Summary of the literature review on strategies and factors to choose an international market

Author	Year	Concept	Contribution
Forsyth;	1972	Factors of attractiveness	Illustrating the effect of market potential (size and growth) and the associated

Weinstein; Khoury; Choi, Tschoegl & Yu; Terpstra &Yu	, 1979 , 1986 , 1988 , 1997	of markets to entrepreneurs	investment risk
Mulligan & Reeve; Gouch; Timmermans	1983 , 1984 , 1986	The factor of financial and socio-economic incentives	Discussing efficient procedures of credit institutions (e.g., banks) to provide funds for entrepreneurial firms. Socio-economic conditions; level of competition, access to markets, access to capital, availability of information about the local tolerance degree, existing supporting networks, and the niche concentration
Oviatt & Mcdougall	1994 , 1995 , 1997	Location decision factors related to technological improvements and development of information technology	- Specialized production and client adoption are more viable for small-scale operations due to improvements in manufacturing technologies. - Advanced transportation offers more reliable, frequent, and cheaper means of movements between countries and continents cutting the cost required to move people and goods. - Easier data accessibility and collection as well as simplified data analysis approaches and interpretation help entrepreneurs identify new opportunities and circumstances that in enabling them to carry on planning & managing international activities from the time of their venture's foundation
Oviatt & McDougall; Madsen & Servais; Rialp et al.	1995 , 1997 , 2005	Internal factors that derive going global	Individual characteristics of the entrepreneur: previous international and business experience, academic training, ambition and motivation levels, risk perception, global vision, leadership and also personal relationship
Jones	1999	Influence of integration in global markets	Manufacturing and services sectors are improving their cross borders networks and links through creative procedures of global supply and distribution
McAuley	1999	Location decision factors	More specialized production and

		related to changes in market conditions	operations strategies to serve specific niches in the international markets that have deficiencies and depend on their distinctive competencies to produce innovative and distinguished products that can be sold worldwide
Andersson	2000	Improving the entrepreneurs' interaction to different cultures	Acquiring skills and more international education and experiences to better understanding of the needs
Rekers and van Kempen; Ram et al.; Schutjens and Stam; Wennekers et al.	2000, 2002, 2003, 2008, 2009	Important factors	Importance of available business opportunities and demographic characteristics, built environment with its local policies and supporting regimes, the increased tendency towards self-employment, the increased outsourcing of business activities by large firms, the rise of internet commerce, and the growing flexibility of labor contracts
Gorzelak	2001	Shift in the twenty-first century towards knowledge-based economies where the markets demand a creative and complex workforce	Critical factors of business attractiveness are categorized into two distinctive groups: - Factors related to resource-based economy labor force, resources, premises, bulk transportation, and energy resources. - Factors related to knowledge-based economy qualification, research and development centers, local supplies, reliable infrastructure, and good living conditions.
Dimitratos & Plakoyiannaki; Zahra et al.	2003, 2005	Components of IEC	Six interrelated organizational culture dimensions: international market orientation, international learning orientation, international innovation propensity, international risk attitude, international networking orientation, and international motivation
Dimitratos & Jones; Zahra, Korri, & Yu	2005	International entrepreneurial culture (IEC)	Considers the international entrepreneurial activities of the firm to identify and pursue opportunities abroad
Rialp et al	2005	Most common deriving factors	New development of market conditions in many sectors of economic activities,

			technological revolutions in production, transportation, communication, etc., global networks and alliances' prosperity, and the growing number of skilled people with entrepreneurial orientation
Struzycki	2006	Governmental and legal support	Increased number and quality of services and resources (e.g., developed land, real estate, etc.), as well as intellectual resources (skills, knowledge, and qualifications of local community members) introduce ambitious plans to provide the basis for creating optimum features for investors, optimizing the use of the limited financial resources and assisting businesses to secure financing from external sources, and better adaptation to environmental changes (arising opportunities or threats), and conducting promotional activities
Hui Tseng, Tansuhaj, Hallagan, & McCullough	2007	Importance of R&D	Research and development intensity is highly important for firms' expansion behavior across borders
Huang, Zhang, Zhao, & Varun	2008	Attractiveness of developing countries	Offer cost savings in the form of lower factory wages along with other attractive business environmental factors such as favorable exchange rates, a significant amount of unskilled labor, and favorable foreign trade policies
Jarczewski	2008	Simplify the process of starting and running business activities	Preparation of real estate with the provision of the physical plan, technical infrastructure and accessible roads, real-estate tax exemptions, and attracting large investor that would most likely promote the goodwill and pro-investment image of the location and consequently accelerate the influx of other businesses
Rialp-Criado, Galvan-Sanchez, & Suarez-Ortega	2010	Small firm to go global from the stage of outset	Derived by internal and external key factors

Beckers & Kloosterman	2011	Several factors the founders of migrant businesses are motivated by to locate their ventures in specific neighborhoods	Knowledge and available information about rules and regulations of the region that are related to a particular line of business have significant importance to obtain the necessary start-up and social capitals, including providing linkages to local suppliers, customers and labors
Cuervo-Cazurra	2011	Most influential internal factor	Decision makers depend heavily on the accumulated amount and type of knowledge the entrepreneurial firm possesses including knowledge about how to manage increased complexity and diversity in international markets, knowledge of the foreign markets, clients, and competitors, and knowledge of foreign government institutional frameworks, rules, norms, and values
Analiza	2012	Governmental and legal support	Locations covered by legal protection, attracting factors such as information and promotional support, grants and subsidies, the advice of business environment institutions, the use of exemption and tax benefits, and assistance in financing as well as in adjusting the profile of requirements to operate in the location
Cannone, Costantino, Pisoni, & Onetti	2012	Some other external factors derive internationalization	Internalization of transactions, an alternative governance structure, the development of the foreign location advantage, and a unique resource control
Kimelberg & Williams	2013	Factors that are used as a basis for location decisions of enterprises in general	Two broad categories (1) studies to measure the influence of a specific factor or a set of factors on firm location decisions, such as analyzing the impact of taxes and incentives, and (2) studies that explain the decision process for a specific business or industry, e.g., the location decision process of biotechnology firms

Table (2.5) Summary of the literature review on strategies to choose an international location

Author	Year	Methodology	Contribution
Yang & Lee	1997	Analytical Hierarchy Process (AHP)	1-Problem decomposition into elements. 2- Comparative analysis: the importance of elements at each level is measured by a procedure of pairwise comparison where each element is prioritized using a rating scale. 3- Synthesis of priorities: priority weight of elements at each level is computed using eigenvector or least square analysis. 4- Location factors: - quantitative: measured in numerical values - qualitative: subjective judgment is adopted
Carlson	2000	(1) Surveys of companies	Ask key respondents about factors led to their location decision, problems: stating of variables that are not quantified, and adopting open-ended questions leading to unintentionally neglected factors
Carlson	2000	(2) Statistical models	Explore some of the factors influenced the selection of a specific location
Beim & Levesque	2003	Multiple Criteria Decision Analysis (MCDA)	1- Selecting a foreign country for new business venturing from the point of view of an entrepreneur. 2- The entrepreneur develop a hierarchy of criteria to assess the countries under consideration under desired criteria 3- Avoid pitfalls of redundancy, lack of independence and complexity. 4- Measurements used best described by categorical labels, not by numerical scores.

2.4 Cluster analysis literature review

Cluster analysis refers to various mathematical methods that are used to determine homogenous groups of objects known as clusters in a set of data (Romesburg, 2004). The objects

in each cluster share many characteristics and have similarities in common, while at the same time they are very dissimilar to objects in other clusters (Springer & Heidelberg, 2011).

There are various methods and algorithms by which the clustering analysis can be applied to perform the data classification (Jain & Dubes, 1988). Some of the most commonly used algorithmic options include:

1. Hierarchical clustering: it is one of the intrinsic genus approaches of classification. This type of clustering includes both agglomerative hierarchical classification and divisive hierarchical classification. In agglomerative hierarchical clustering, each object is placed in its own cluster followed by gradual merging of these atomic clusters into larger and larger clusters until all objects can be combined into one large single cluster. On the other hand, the process of divisive hierarchical clustering starts with having all objects in one cluster that will be subdivided into smaller pieces.
2. Partitional clustering: it is another intrinsic genus approaches of classification that also includes agglomerative classification; small clusters are joined together to form a single partition and divisive classification that is carried out by fragmenting a single all-inclusive cluster.
3. Serial and simultaneous clustering: the patterns are handled one by one in the serial classification, whereas, in simultaneous classification the entire set of patterns is operated at the same time.
4. Monothetic and polythetic clustering: in monothetic clustering the features are used one by one, while all the features are used at once in polythetic clustering.

For a variety of research goals, scholars and researchers from all fields need to find out which objects are similar or dissimilar in a set of data. A prominent research goal for which the cluster analysis is favorably used is building up data classification (Romesburg, 2004). Therefore, applications of cluster analysis are useful in all professions. Cluster analysis can satisfactorily fulfill different purposes in science, planning, management, as well as many other research fields.

The decision making process as a genuine component of planning and management activities can also benefit from the applications of cluster analysis in which the available alternative decisions or plans represent the objects of the cluster analysis whereas the attributes describe the features or the expected outcomes of the alternatives. The identified clusters of similar alternatives would then reduce the decision problem into only two phases selecting the cluster that best achieves the planning objective, and then selecting the best alternatives within the best cluster (Romesburg, 2004).

Several clustering methods are used to perform the cluster analysis, particularly to reduce the size of the resemblance matrix. The clusters that are generated through performing clustering methods are comprised of a number of points. In a multi dimensional space, each of these points is usually represented by a vector of values. In order to decide which clusters to be merged or split, a combination of two factors is used to obtain a measure of similarity/dissimilarity measure between clusters (Anandan, 2013);

1. Distance Metric: used to find the distance between two points (represented by vectors), e.g. the Euclidean distance.

   The Euclidean distance between two points that are represented by the vectors $p = (p_1, p_2, \cdots, p_n)$ and $q = (q_1, q_2, \cdots, q_n)$ are given by

$$d(p,q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2}$$

2. Linkage Criteria: used to find the distance between two clusters. This distance is calculated by deciding on how to use the points of each cluster. A particular linkage criterion should be selected and used in conjunction with a distance metric to find the distance between the clusters.

Some of the commonly used linkage criteria include the single linkage-clustering method (SLINK), the complete linkage-clustering method (CLINK) and the average linkage-clustering (ALC) or the unweighted pair-group method using arithmetic averages (UPGMA).

The suggested model in this research demands adopting a clustering method to obtain clusters in which the addition of an entity to a cluster must not require that the entity is highly similar to any member of that cluster, i.e., preventing the chaining reaction (formation of clusters that can tend to resemble long chains). The Complete Linkage Clustering (CLINK) or the Average Linkage Clustering (ALC) are the most appropriate clustering algorithms to satisfy this requirement. However, implementing the CLINK analysis exceeds any of the other hierarchical clustering approaches in fulfilling this requirement and other preferred characteristics such as generating small and tightly bound clusters and for the tendency to prevent merging two clusters for only the high level of similarity between two members when the remaining members are dissimilar. More details on the different types of the hierarchical clustering algorithms are given in the following section.

2.4.1 Similarity based clustering

McAuley, based on the Jaccard similarity coefficient, introduced an early definition of the similarity coefficient-based clustering concept in 1972. In McAuley's definition, the

similarity coefficient between any two objects represents the ratio of the number of attributes that belong to the two objects to the sum of the number of attributes that belong to either or both of the objects. In 1973, Carrie generalized the same similarity coefficient approach to become the value that is calculated for each pair of attributes instead of the objects (Wang and Roze, 1984).

According to Gupta and Seifoddini (1990), the Similarity Coefficient Method (SCM) outperforms other clustering approaches through providing various advantages when it is implemented, including the following:

- It is simpler and easier to be used with computer applications
- It is more flexible in incorporating additional quantitative and subjective information into the formation process of machine cells.
- It intrinsically determines the level of similarity (the threshold value) by which two or groups of machines are allowed to form for each iteration of a given set of data in problems.
- It permits consideration of additional constraints for the final selection of a solution through generating a set of alternative solutions.

On the other hand, the SCM's major drawback of not accounting for many important variables in the Jaccard similarity coefficient stimulated further research work on the subject. As a result, a new algorithm was developed based on the similarity coefficient method (SCM) for the purpose of grouping the machines into machine cells by using complete linkage clustering (CLINK) with the incorporation of various important production parameters such as part type production, volume, routing sequence, and unit operation time (Gupta and Seifoddini, 1990).

In 1998, Nair and Narendran suggested another new similarity coefficient, in which the similarity coefficient is calculated based on the sequence of parts and yielding a higher quality clustering. A year later, Nair and Narendran (1999) prepared a paper to discuss another similarity coefficient method that takes into account additional similarity coefficients' calculating information such as production sequence, production volumes, processing times, and the capacity of machines. Furthermore, Table (2.7) includes more of the literature review on similarity based clustering and Figure (2.1) below illustrates the considered and applied similarity coefficient-based clustering and the related similarity measures in this research.

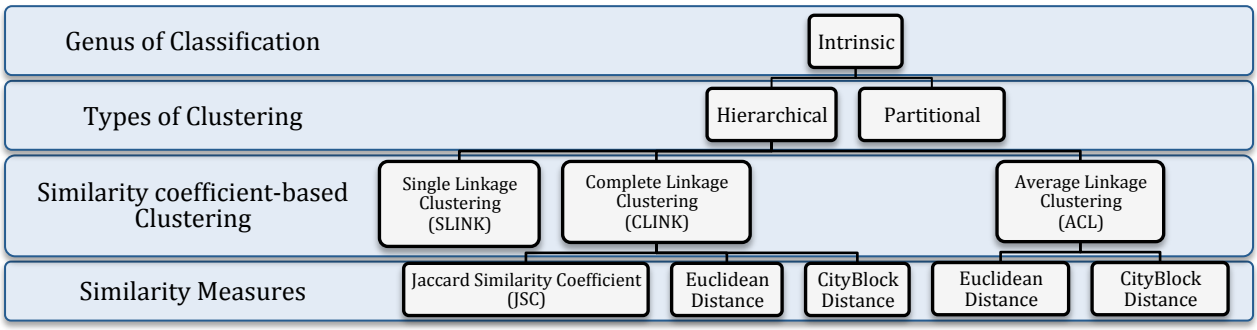

Figure (2.1) Considered similarity coefficient-based clustering and similarity measures

Another interesting clustering method is the rank order-clustering algorithm (ROC). The ROC algorithm can be used in synchronization with a block and slice method in order to form a set of intersecting machine cells and non-intersecting part families. After obtaining this set, a hierarchical clustering method is applied based on a similarity measure among the machine pairs. Chandrasasekharan and Rajagopalan (1986) were also able to introduce a non-hierarchical clustering approach for the concurrent formation of part families and machine cells in 1987. The proposed algorithm begins with a clustering algorithm that is run based on representative seeds. Performing a block diagonalization algorithm then follows the formation of the clusters. The last step is applying a clustering algorithm that is based on ideal seeds to modify the previously

generated clusters. To efficiently identify the required seeds, in 1991 Srinivasan and Narendran explored the issue more and developed a convenient non-hierarchical clustering algorithm.

2.4.2 Methods of similarity coefficient-based clustering

In the machine-part cellular manufacturing, the similarity coefficient-based clustering methods rely on similarity measures in conjunction with clustering algorithms. These methods usually consist of a standard set of the following main steps: (Yin and Yasuda, 2006)

1. Formation of the machine-part incidence matrix, in which rows are for the machines and columns stand for parts. The entries in the matrix are either 0s or 1s depending on the need of a part to be processed on a machine or not. Any entry in the matrix a_{ik} is defined as

$$a_{ik} = \begin{cases} 1 & \text{if part } k \text{ visits machine } i \\ 0 & \text{otherwise} \end{cases},$$

   where i is the machine index ($i = 1, \ldots, M$) for M number of machines and k is the part index ($k = 1, \ldots, P$) for P number of parts.
2. Selection of a similarity coefficient to calculate the similarity values between machine (part) pairs and to create the similarity matrix in which the elements represent the similarity between two machines (parts).
3. Implementing a clustering algorithm to process the values in the similarity matrix to obtain a diagram known as a tree or a dendrogram, which shows the similarities hierarchy among all pairs of machines (parts).
4. Identifying the groups of machines (part families) from the resulting dendrogram and checking all predefined constraints such as the number of cells, cell size, etc.

One of the earliest and most commonly used similarity coefficients to measure the similarity among objects is the Jaccard Similarity Coefficient (JSC) (Wang and Roze, 1984). In

the JSC approach (a machine clustering example is given for simplification purposes), the similarity coefficient is calculated depending on the number of parts visiting each machine. Also, all attributes are set to be binary and therefore the yielded possibilities for each pair of machines are: 1, 1 or 0, 0 or 1, and 0 as indicated in Table (2.6) below.

Table (2.6) Yielded possibilities for the attributes in JSC

		Machine *j*	
		1	0
Machine *i*	1	a	b
	0	c	d

(Saiful Islam & Sarker, 2000)

Where a is the number of parts visiting both machines *i* and *j*, b is the number of parts visiting only machine *i*, c is the number of parts visiting only machine *j*, and d is the number of parts visiting neither machine *i* nor machine *j*.

Then, JSC is calculated by the formula

$$s_{ij} = \frac{a}{a+b+c}, \quad 0 \leqslant s_{ij} \leqslant 1 \qquad \text{(Yin and Yasuda, 2006)}$$

Moreover, the Jaccard similarity coefficient suggests that

- The value range of the similarity coefficient is between 0 and 1,
- The maximum value is obtained when the same parts are processed by both machines, i.e., $b = c = 0$, and
- The minimum value is obtained when none of the parts visit both machines, i.e., $a=0$.

Another similarity measure that is used to measure the similarity between two clusters is the Euclidean distance. The Euclidean distance between two clusters, cluster A that has the mean vector A = (x_{a1}, x_{a2}, …, x_{am}) and cluster B that has the mean vector B = (x_{b1}, x_{b2}, …, x_{bm}) is calculated as

$$d(\mathrm{A}, \mathrm{B}) = \left(\sum(xai - xbi)^2\right)^{1/2}$$ (Salameh, 2000).

The CityBlock distance (Manhattan distance) is also a similarity measure where the distance between two points in the xy-plane is calculated as the distance in x plus the distance in y, which is similar to moving around the buildings in a city (like the city of Manhattan) instead of going straight through.

The CityBlock distance between two points a ∈ cluster A and b ∈ cluster B is calculated as follows: (Zhang and Lu, 2003)

$$d(\mathrm{A}, \mathrm{B}) = \sum_{j=1}^{m} |a_j - b_j|$$, where $j = (1, 2, \ldots, m)$ is the attribute

The CityBlock distance is always greater than or equal to zero. It equals zero for the identical similarity while it is high for the little similarity.

Many methods for data clustering are available and the considered dataset may be grouped in various different fashions depending on the type of clustering method that is used. Therefore, the selection of a particular method depends mainly on the desired output type. Also, selecting the clustering method is most likely affected by several unique characteristics of the chosen method, including the performance of the method with specific data type, the available hardware and software facilities for the selected method, and the size of the dataset the method can handle.

Following are some of the most commonly used data clustering methods along with a brief approach of execution for each of them (illustration of the implementation of algorithms is carried out using machine clustering as an example for simplification purposes).

Single Linkage Clustering (SLINK)

The single linkage-clustering algorithm is the one best-known method of hierarchical clustering that Sneath first developed in 1973. It is also known by the names (minimum method) and (nearest neighbor cluster analysis), characterized by its minimal computational requirements among all the similarity coefficient-based clustering algorithms. At each step in the SLINK algorithm, the two most similar objects that are not yet in the same cluster are joined. In fact, the term single linkage implies the act of joining pairs of clusters by the single shortest link between them (Tamilselvi, Sivasakthi, and Kavitha, 2015).

The distance between two clusters X and Y in the single linkage-clustering (SLINK) is calculated as the distance between the two closest points $x \in X$ and $y \in Y$.

$$d(X,Y) = \min_{x \in X, y \in Y} d(x,y)$$ (Anandan, 2013)

The SLINK algorithm starts with the calculation of similarity coefficients for each pair of machines that is followed by the formation of the similarity matrix. In order to determine the minimum similarity coefficient value through which two machines would be considered similar, the decision maker is required to identify a specific threshold. After setting up the matrix, machines having the highest similarity coefficient are grouped together. Then, the same process is repeated until the maximum value of the similarity coefficient for the unassigned machine to any of the clusters drops below the predefined threshold value or the predefined number of clusters.

In general, the SLINK algorithm is executed in the following standard steps:

1. Set up the similarity matrix by calculating the similarity coefficient for each pair of machines.
2. Determine the groups of machines with the maximum similarity coefficient and put them together.
3. Eliminate the rows corresponding with the machine groups that were grouped together.
4. Add a new row to the matrix for the resulting new machine group and compute the similarity coefficient using the formula $S_{tv} = Max\{S_{mn}\} \quad m \in t \ \& \ n \in v;$ Where t is the new machine group and v is for the other machine groups.
5. Repeat the steps from step 2 to step 4.
6. The algorithm terminates when the number of machine groups that was previously determined is achieved.

Furthermore, the cluster in the SLINK analysis is defined as a group of entities such that every member of the cluster is more similar to at least one member of the same cluster than it is to any member of another cluster.

Adding an entity to a cluster in the single linkage cluster analysis requires that the entity is highly similar to any member of that cluster and due to this procedure, the formed clusters can tend to resemble long chains in multidimensional space. This tendency to chain is considered as a major drawback of the SLINK cluster analysis. A simple example on this feature is a clustering problem that has five entities A, B, C, D, and E, where A is similar to B, which is similar to C, which is similar to D, leading to ABCD would form a cluster. However, the entities A and D might exhibit a relative dissimilarity to each other and each of them might show a higher

similarity to the entity E than to each other. In fact this chaining phenomenon have induced the rejection of the SLINK analysis as a preferable clustering procedure (Blashfield, 1976).

Complete linkage Clustering (CLINK)

The complete linkage-clustering algorithm is also one of the hierarchical clustering methods. It is also known by other different names, (maximum method) and (furthest neighbor cluster analysis). In this algorithm, the least similar pair between two clusters is used to determine the inter-cluster similarity, i.e., the member of every cluster is more like the furthest member of its own cluster than the furthest item in any other cluster (Tamilselvi, Sivasakthi, and Kavitha, 2015).

In the (CLINK) method, the distance between two clusters X and Y is computed as the maximum distance between any two points $x \in X$ and $y \in Y$ in the two clusters.

$$d(X,Y) = \max_{x \in X, y \in Y} d(x,y)$$ (Anandan, 2013)

In the complete linkage clustering, the clusters are small and tightly bound, with the advantage of preventing the merge of two clusters together for only the high level of similarity between two members when the remaining members are dissimilar. Therefore, the cluster in the CLINK analysis can be defined as a group of entities in which each member is more similar to all the other members within the same cluster than it is to all members of any other cluster. Such properties make the complete linkage method able to overcome the tendency to chain issue of the single linkage method.

On the other hand, an entity in the complete linkage method cannot join a cluster until it obtains a given similarity level with all members of a cluster which leads to lowering the probability of obtaining a new member as the cluster size increases. In the multidimensional

space, this means that as the size of a cluster increases, the effective distance between the cluster and nonmember also increases creating what is known as the CLINK's space-diluting feature (Blashfield, 1976).

Average linkage Clustering (ALC)

Unlike the single linkage method that is based on the maximum similarity, or the complete linkage method in which the minimum similarity is the basis, the average linkage-clustering algorithm considers the average value of the pair wise within a cluster (Tamilselvi, Sivasakthi, and Kavitha, 2015).

The average linkage clustering (which some scholars also call it the Unweighted Pair Group Method using Arithmetic Mean (UPGMA)) is considered as a compromise between the chaining tendency of single linkage clustering and the space-diluting tendency of complete linkage clustering (Blashfield, 1976).

In this algorithm, and due to the fact that all objects in a cluster contribute to the inter-cluster similarity, each object is more similar to every other member of its own cluster than to the objects in any other cluster on average and the distance between two clusters is calculated by the average of the distances between all the points in the two clusters.

$$d\ (X,Y) = \frac{1}{|X|\ .\ |Y|} \sum_{x \in X} \sum_{y \in Y} d(x.y)$$ (Anandan, 2013)

where x is any point in the cluster X and y is any point in the other cluster Y.

Standard steps for the ALC algorithm are:

1. Set up the similarity matrix by computing the similarity coefficients for each pair of machines.
2. Allocate in one group all the machine groups of the highest similarity coefficient.

3. Eliminate the rows corresponding with the machine groups that have been grouped together.
4. Add a new row to the resulting matrix for the new machine group and compute the similarity coefficients using the formula to calculate the similarity between the machine groups in the ALC algorithm

$$S_{tv} = \frac{\sum_{m \in t} \sum_{n \in v} S_{ij}}{N_t * N_v}$$

where t is the new machine group and v is for the other machine groups.

5. Repeat the steps from step 2 to step 4.
6. The algorithm terminates when the number of machine groups that was previously determined is reached.

The cluster in the average linkage cluster analysis is defined as a group of entities in which each member has a greater mean similarity with all members of the same cluster than it does with all members of any other cluster (Blashfield, 1976).

The proposed model in this research is basically derived from the clustering analysis approach utilized to study the formation of clusters of machine cells visited by part families based on specified attributes of the parts. Similarly, a similarity coefficient-based clustering algorithm is implemented in this research, namely the complete linkage-clustering method (CLINK), to create clusters of similar countries that have the potential to offer the best locations to start up entrepreneurial ventures with the consideration of factors that are appealing to entrepreneurs. For entrepreneurs, in general, it is more desirable to have more distinct groups of alternate locations (countries) in which the alternatives within each group of locations (countries) are more similar to each other than to the locations (countries) in the other groups. This approach

provides the entrepreneurs with more flexibility in choosing the location for their ventures and businesses from the identified alternatives in the same group with less overlapping between the groups that are distinguished in the level of their entrepreneurial attractiveness.

Table (2.7) Summary of the literature review on cluster analysis

Author	Year	Concept	Contribution
Springer & Heidelberg	2011	Similarity and dissimilarity in cluster analysis	Objects in each cluster share many characteristics and have similarities in common, while at the same time they are very dissimilar to objects in other clusters
Romesburg	2004	Usefulness of cluster analysis	- Cluster analysis is very useful and satisfactory in building up data classification - The available alternative decisions or represent the objects of the cluster analysis whereas the attributes describe the features or the expected outcomes of the alternatives
McAuley	1972	Definition of the similarity coefficient-based clustering	Similarity coefficient between any two objects represents the ratio of the number of attributes that belong to the two objects to the sum of the number of attributes that belong to either or both of the objects
Wang & Roze	1984	Definition of the similarity coefficient-based clustering	The similarity coefficient approach is generalized to become the value that is calculated for each pair of attributes instead of the objects
Gupta and Seifoddini	1990	Advantages of implementing similarity coefficient based clustering	- It is simpler and easier to be used with the computer applications - It is more flexible in incorporating additional quantitative and subjective information into the formation process of machine cells - It intrinsically determines the level of similarity (the threshold value) by which two or groups of machines are allowed to form for each iteration of a given set of data in problems

			- It permits consideration of additional constraints for the final selection of a solution through generating a set of alternative solutions
Wang & Roze	1984	Jaccard Similarity Coefficient (JSC)	- Similarity coefficient is calculated depending on the number of parts visiting each machine - All attributes are set to be binary and therefore the yielded possibilities for each pair of machines are: 1-1, 1-0, 0-1, and 0-0
Tamilselvi, Sivasakthi & Kavitha	2015	Single linkage-clustering (SLINK)	- minimal computational requirements - At each step: the two most similar objects that are not yet in the same cluster are joined - Joining pairs of clusters by the single shortest link - Alternatives having the highest similarity coefficient are grouped together
Tamilselvi, Sivasakthi & Kavitha	2015	Average Linkage Clustering (ALC)	- Considers the average value of the pair wise within a cluster - Each object is more similar to every other member of its own cluster than to the objects in any other cluster on average
Tamilselvi, Sivasakthi & Kavitha	2015	Complete Linkage-Clustering (CLINK)	- The least similar pair between two clusters is used to determine the inter-cluster similarity - The member of every cluster is more like the furthest member of its own cluster than the furthest item in any other cluster - Clusters are small and tightly bound - Prevents the merge of two clusters together for only the high level of similarity between two members when the remaining members are dissimilar.

CHAPTER THREE

Identifying the entrepreneurial location decision factors

As stated in the previous chapters, choosing to enter a foreign market might be one of the most critical strategic decisions a firm has to encounter. Moreover, the consequences of the location decision have more effects and larger impacts when the firm is small in size and entrepreneurial in nature due to the limited resources available.

Like any other decision, the decision-making process of determining the best location for small entrepreneurial firms features the need to identify potential alternatives or options that are must be evaluated by the decision maker in order to specify the best alternative. In the location decision problem, the potential alternatives are the possible sites to locate the firm that have to be evaluated by the entrepreneur/founder and then to choose the best from among them.

However, identifying the best location for a facility is not an easy task and particularly for a small enterprise, because personal characteristics of the founder/entrepreneur usually have a great influence on the decision-making process. In fact, all strategic decisions within the small firms are influenced by the entrepreneurial characteristics of their founders. Therefore, it is essential to consider the entrepreneurial behavior effects in the decision-making process of small firms. In small firms, it is expected that the rationality trait is decreased in proportion to the higher impact of the entrepreneur's personality. The optimistic nature of entrepreneurs also may cause their decisions to be based on subjective factors.

Similarly, choosing best-fit locations for the facility is greatly affected by the individual personality traits and cognitive biases of the entrepreneur, including the need for achievement, the locus of control, the optimum risk propensity, and innovativeness. This, in addition to the

complex nature of the decision location problem, increases the relevance of the factors provided by the entrepreneur, who in most cases is the primary decision maker in the small enterprise, to make a better locational judgment.

In general, there is no single valid solution for all location decision problems and therefore choosing an optimal location for the facility demands careful analysis of all critical subjective factors to assess the various potential locations.

Furthermore, the previous chapters have shown that researchers around the world have carried out the mission to develop various algorithms and techniques with the aim to provide the decision makers with reliable tools to promote their location decision approaches. Through these algorithms, the facility location problem is addressed from different angles. The application of each of these algorithms most likely leads to identifying alternatives as best choices that are unique and different for each algorithm based on its own perspective, and the best choices generated by one algorithm do not necessarily have to be favored by the other approaches.

Moreover, the empirical implementation of the algorithms mainly depends on comparing the different alternatives in accordance with a set of pre-defined decisive factors. The set of these factors should be provided by the decision maker in a comprehensive context that takes into account all different aspects of the location decision case, because failing to include one or more of the substantial factors may result in developing ineffective or misleading decisions upon the best location of the firm.

The suggested model to the facility location problem in this research investigates the similarities and dissimilarities of alternate sites that have the potential to locate the small firms within and classifies them into distinctive groups based on a set of decision-making factors.

In this chapter, the most critical judgmental factors are derived basically from the revised related literature. To properly use these factors in this research, they need to be broken into the most relevant sub-factors or indicators for which numerical data are available in the global indices. One of the most important global indices that contain comprehensive data about development in countries around the globe is the World Bank's developmental indicators index.

Depending on the World Bank's index, all associated sub-factors are defined throughout this chapter and then they are used in the subsequent chapters to represent the core of the required determinants that in return, are employed to efficiently classify the groups of locations and assure valid results when conducting the location decision algorithm.

Moreover, the identified decision-making factors in this chapter include the factors existing in potential locations that are most attractive to entrepreneurial firms or, if different, factors that local governments strive to implement into regions under their authorities to offer a favorable economic climate for new businesses.

Based on distinctive criteria of attraction to entrepreneurs, the most likely location decision factors that should be considered in choosing the best-fit location for small and medium entrepreneurial enterprises can be specified as follows:

3.1 Factors related to business start-up cost and procedure

A favorable legal system regarding incorporation, organizational, and publicly held status of a small venture has important implications for its behavior, growth and success. Therefore, decision makers need to clearly study and understand existing corporate and securities laws in considered sites to locate the entrepreneurial facility. On the other hand, special consideration of small and medium enterprises, such as specific exemptions from regulations, modified

compliance procedures, reduced penalties for violation of regulations, and specialized programs to assist small and medium enterprises in compliance with regulations, should be embedded in the policymaking process for the region to strengthen its appeal to new business. Some of the most important attributes that most likely influence the choice of localizing entrepreneurial firm globally are:

3.1.1 Cost of business start-up procedures (% of GNI per capita)

This factor consists of the necessary expenses the business is required to spend in order to acquire a sound legal structure at the establishment stage, including registration fees and permits and licenses charges, etc., for the business to be qualified to start its operations.

3.1.2 Start-up procedures to register a business (total number)

This factor contains all related procedures of ownership, size, and type of business that are required to start up the business, such as interactions to obtain necessary permits and licenses and to complete all inscriptions, verifications, and notifications to start operations.

3.1.3 Time required to start a business (days)

It is the number of calendar days needed to complete all needed procedures to legally start operating the business. The fastest procedure is considered even if additional costs are required to speed up one or several of the procedures.

3.1.4 Patent applications, nonresidents (total number)

These are worldwide patent applications filed through the Patent Cooperation Treaty procedure or with a national patent office for exclusive rights for an invention, a product, or process that introduces a new way of doing something or offers a new technical solution to a

problem. The obtained patent provides protection for the patented material to its owner for a specified period of time, generally 20 years.

3.1.5 Trademark applications, direct nonresident (total number)

Filed trademark applications are those applications to register a trademark with a national or regional Intellectual Property (IP) office. A trademark is a distinctive sign that identifies certain goods or services that are produced or provided by a specific person or enterprise. The importance of a trademark is to provide protection of the mark to its owner by ensuring the exclusive rights to use it in identifying goods or services, or to authorize others to use it in return for named payment. Protection periods vary; however, a trademark can be renewed indefinitely beyond the time limit on payment of additional fees. Specifically, direct nonresident trademark applications are those that are filed by applicants from abroad directly at a national IP office.

3.1.6 Charges for the use of intellectual property, payments (current US$)

These are the payments and receipts between residents and nonresidents for the authorized use of proprietary rights (patents, trademarks, copyrights, industrial processes and designs including trade secrets, and franchises) and for the use, through licensing agreements, of produced originals or prototypes (such as for live performances and television, cable, or satellite broadcast).

3.2 Factors related to financing small and medium enterprises

Financial resources are the life-blood for any business and specifically for entrepreneurial firms, small or medium, sufficient funds are vital to meet daily expenses and payments. Typically, the main sources of funding for the business revenues from the business operations come from investments of the owner, a partner, or a venture capitalist, and loans from

individuals or financial institutions. Thus, the availability of specialized financial policies and incentives is a significant determinant for choosing a location to start entrepreneurial firms. Some of the critical factors to consider in selecting a globally appealing location include:

3.2.1 Firms using banks to finance investment (% of firms)

This gives an indication of the percentage of firms using available services of banks to finance their investments.

3.2.2 Lending interest rate (%)

The lending interest rate is the defined rate by banks that usually meets the short- and medium-term financing needs of the private sector. This rate is normally differentiated according to the creditworthiness of borrowers and the objectives of the financing. Terms and conditions of these rates differ by country.

3.2.3 Foreign direct investment, net (current US$)

These are the net inflows of investment to acquire a lasting management interest (10 percent or more of voting stock) in an enterprise operating in an economy other than that of the investor. The foreign direct investment is the sum of equity capital, reinvestment of earnings, other long-term capital, and short-term capital as shown in the balance of payments.

3.3 Factors related to tax rates and structure

The personal income taxes along with capital gains taxes and payroll taxes result in leaving the individual entrepreneurs with less expendable capital, i.e., the higher the tax rate, the more capital is taken from the business and given to the government. Higher tax rates means less money to reinvest in the business, leading to less job creation. Therefore, it is critical to assess

the effect of tax rates and structures when exploring the possibility of choosing a country to locate the entrepreneurial firm.

3.3.1 Total tax rate (% of commercial profits)

This is a measurement of the amount of taxes and mandatory contributions payable by businesses after accounting for allowable deductions and exemptions as a share of commercial profits. In this indicator, taxes withheld (e.g., personal income tax or collected and remitted to tax authorities (such as value added taxes, sales taxes, or goods and service taxes) are excluded.

3.3.2 Profit tax (% of commercial profits)

This is the amount of taxes paid by the business on its profits.

3.3.3 Taxes on goods and services (% value added of industry and services)

The taxes on goods and services include general sales and turnover on value added taxes, selective excises on goods, selective taxes on services, taxes on the use of goods or property, taxes on extraction and production of minerals, and profits of fiscal monopolies.

3.4 Factors related to governmental regulations and policies

Governments establish many rules and regulations that organize and control the business environment of the country. As a result, businesses need to plan their operations' structure to comply with the governmental regulations. Furthermore, economical policies and market regulations have a significant impact on the competitiveness and profitability of the business and therefore, choosing a country to locate the entrepreneurial firm is influenced heavily by the type of the governmental rules and policies applied in that country.

3.4.1 Exports of goods and services (current US$)

This category comprises all transactions between residents of a country and the rest of the world, involving the change of ownership from residents to nonresidents of general merchandise, the net exports of goods under merchanting, nonmonetary gold, and services.

3.4.2 Trade in services (% of GDP)

Trade in services is the sum of services exports and imports divided by the value of GDP.

3.4.3 Net official development assistance and official aid received (current US$)

Net official development consists of disbursements of loans made on concessional terms (net of repayments of principal) and grants by official agencies of the members of the development assistance committee (DAC), by multilateral institutions, and by non-DAC countries to promote economic developments and welfare in countries and territories in the DAC list of ODA recipients. It includes loans with a grant element of at least 25 percent (calculated at a rate of discount of 10 percent).

Net official aid is the aid flow (net of repayments) from official donors to countries and territories.

3.5 Factors related to labor and skills

Access to skilled labor is critical to entrepreneurial firms. Innovative businesses require specific skills and experience and skilled labor can actually contribute to innovation and growth activities of the firm by generating new knowledge, developing incremental innovations, supporting firms in identifying business opportunities, helping firms to adopt to changing environments, and generating spillovers and transfer of advanced knowledge. From their side, public policies can facilitate new firms' access to skilled labor via strengthening education about

innovation, promoting the innovation culture, and elevating investment rate in research and development activities.

The availability of a wider spectrum of skills in a labor market within a country significantly encourages entrepreneurial companies to locate their facilities in that country.

3.5.1 Labor force with tertiary education (% of total)

It is the share of the total labor force that attained or completed tertiary education as the highest level of education.

3.5.2 Secondary education, vocational pupils (total number)

This factor is the total number of students enrolled in technical/vocational programs at public and private secondary education institutions.

3.5.3 Government expenditure on education, total (% of government expenditure)

General expenditure of the government on education (current, capital, and transfers) is expressed as a percentage of the total general government expenditure on all sectors (including health, education, social services, etc.). This also includes expenditure funded by transfers from international sources to the government. General government refers to local, regional and central governments.

3.5.4 Wage and salaried workers, total (% of total employed)

Wage and salaried workers (employees) are the workers who hold the type of jobs defined as "paid employment jobs," where the incumbents hold explicit (written or oral) or implicit employment contracts that give them a basic remuneration that is not directly dependent upon the revenue of the unit for which they work.

3.5.5 Unemployment, total (% of total labor force)

Unemployment is the share of labor force that is without work but available for and seeking employment.

3.6 Factors related to infrastructure

The availability of a good infrastructure that may take different forms and functions is important for entrepreneurial activities. One form of infrastructure that attracts the attention of entrepreneurs to situate their companies in a country is a high standard of physical infrastructure, including roads, rails, and water routes, which is required for trade and industrial growth.

3.6.1 Investment in energy with private participation (current US$)

This covers infrastructure projects in energy (electricity and natural gas transmission and distribution) that have reached financial closure and directly or indirectly serve the public. The included types of projects are operations and management contracts, operation and management contracts with major capital expenditure, greenfield projects (in which a private entity or public-private joint venture builds and operates a new facility), and divestitures. Investment commitments are the sum of investments in facilities and investments and investments in government assets. Investments in facilities are the resources the project company commits to invest during the contract period either in new facilities or in the expansion and modernization of existing facilities. Investments in government assets are the resources the project company spends on acquiring government assets such as state-owned enterprises, rights to provide services in a specific area, or the use of specific radio spectra.

3.6.2 Investment in telecoms with private participation (current US$)

This covers infrastructure projects in telecommunication that have reached financial closure and directly or indirectly serve the public. The types of projects included are operations and management contracts, operation and management contracts with major capital expenditure, greenfield projects (in which a private entity or public-private joint venture builds and operates a new facility), and divestitures. Investment commitments are the sum of investments in facilities and investments and investments in government assets. Investments in facilities are the resources the project company commits to invest during the contract period either in new facilities or in expansion and modernization of existing facilities. Investments in government assets are the resources the project company spends on acquiring government assets such as state-owned enterprises, rights to provide services in a specific area, or the use of specific radio spectrums.

3.6.3 Investment in transport with private participation (current US$)

This covers infrastructure projects in transport that have reached financial closure and directly or indirectly serve the public. The included types of projects are operations and management contracts, operation and management contracts with major capital expenditure, greenfield projects (in which a private entity or public-private joint venture builds and operates a new facility), and divestitures. Investment commitments are the sum of investments in facilities and investments and investments in government assets. Investments in facilities are the resources the project company commits to invest during the contract period either in new facilities or in expansion and modernization of existing facilities. Investments in government assets are the resources the project company spends on acquiring government assets such as state-owned enterprises, rights to provide services in a specific area, or the use of specific radio spectrums.

3.6.4 Investment in water and sanitation with private participation (current US$)

This covers infrastructure projects in water and sanitation that have reached financial closure and directly or indirectly serve the public. The included types of projects are operations and management contracts, operation and management contracts with major capital expenditure, greenfield projects (in which a private entity or public-private joint venture builds and operates a new facility), and divestitures. Investment commitments are the sum of investments in facilities and investments and investments in government assets. Investments in facilities are the resources the project company commits to invest during the contract period either in new facilities or in expansion and modernization of existing facilities. Investments in government assets are the resources the project company spends on acquiring government assets such as state-owned enterprises, rights to provide services in a specific area, or the use of specific radio spectrums.

3.7 Factors related to technology advancement

Another form of infrastructure is the knowledge infrastructure, which is a crucial feature specifically for knowledge- and technology-based ventures. The existence of research facilities such as universities is also important, since they represent ideal incubators to assist entrepreneurs to benefit from information and knowledge spillovers. Provision and quality of the knowledge infrastructure is a key driver for firm foundation and subsequent economic growth.

3.7.1 High-technology exports (current US$)

High-technology exports are products with high research and development intensity, such as in aerospace, computers, pharmaceuticals, scientific instruments, and electrical machinery.

3.7.2 Internet users (per 100 people)

Internet users are individuals who have used the internet (from any location) in the last 12 months. Internet can be used via a computer, mobile phone, personal digital assistant, games machine, digital TV, etc.

3.7.3 Fixed broadband subscriptions (per 100 people)

It refers to the fixed subscriptions to high-speed access to the public Internet (a TCP/IP connection), at downstream speeds or equal to, or greater than 256kbit/s. Internet subscriptions include cable modem, DSL, fiber-to-the-home/building, other fixed (wired)-broadband subscriptions, satellite broadband and terrestrial fixed wireless broadband. This total is measured irrespective of the method of payment. Moreover, the Internet users are the individuals who have used the Internet (from any location) in the last 12 months.

3.7.4 Research and development expenditure (% of GDP)

The expenditures for research and development are current and capital expenditures (both public and private) on creative work undertaken systematically to increase knowledge, including knowledge of humanity, culture, and society, and the use of knowledge for new applications. The research and development (R&D) includes basic research, applied research, and experimental development.

3.7.5 Researchers in research and development (R&D) (per million people)

These are the professionals engaged in the conception or creation of new knowledge, products, processes, methods, or systems and in the management of the projects concerned. Postgraduate Ph.D. students engaged in research and development are included.

3.7.6 Manufacturing, value added (current US$)

Manufacturing refers to industries and the value added is the net output of a sector after adding up all outputs and subtracting intermediate inputs. It is calculated without considering deductions for depreciation of fabricated assets or depletion and degradation of natural resources. The value added origin is determined by the International Standard Industrial Classification (ISIC).

3.8 Factors related to competition

In today's increasingly open and integrated global economy, competitiveness both domestically and internationally has become a prominent concern. Rapid changes in the global business environment, including trade liberalization, technological development, and governmental policies associated with globalization have simplified the entry of firms to different geographic markets that, in turn, increased the competitiveness level of firms around the world. Although the globalization phenomenon has considerably enhanced the market opportunities of start-up firms, at the same time it also contributed heavily to increasing the amount of competition faced by such firms. It is important for start-up businesses to take into account to a far extent the intensity of the competitive atmosphere when selecting a country in which to locate their facilities.

3.8.1 Listed domestic companies (total number)

They are the domestically incorporated companies listed on the country's stock exchanges at the end of the year. Investment companies, mutual funds, and other collective investment vehicles are not included in this factor.

CHAPTER FOUR

Model Description and Methodology

To achieve the core objective of this research, that is to assign countries into homogenous groups based on their level of attractiveness to entrepreneurs, an efficient clustering method has to be applied. Building up these homogenous groups requires the identification of the decision-making factors upon which similarities and dissimilarities of countries to form clusters are specified.

In the previous chapter, the most critical factors attracting entrepreneurial small and medium ventures to a location have been identified. This task has been carried out through first reviewing the literature discussing why and what attracts entrepreneurial activities to a site. Then, these publications were carefully examined in order to extract important attributes characterizing entrepreneurship-appealing locations. Finally, the yielded factors that are adopted in the model of this research are those that frequently appeared in the related literature and researches or those that are emphasized by experienced and specialized scholars.

Prior to applying the model used in this research, data denoting the location-decision factors have to be collected. It is important that these data are represented with numerical figures in order to provide the model with a mean to measure the considered factors.

4.1 Data collection and setup

In order to better study the decision-making factors and utilize them to assist entrepreneurs to choose an optimal location for establishing their start-up entrepreneurial facilities, numerical data influencing the effects of location decision factors have to be collected

from reliable and trusted entrepreneurship indices or global database sources to assure higher accuracy of data.

There are many indices that convey numerical data that measure the effects of various attributes considered in deciding upon locations where entrepreneurial activities can be started up. The numerical data intended to be collected for the purposes of this research are mainly derived from the World Bank's database. Database from the World Bank surpasses its counterparts based on several unique features, such as being one of the most authentic database sources, as well as the availability of many of the desired numerical data for considered decision-making factors.

However, data collection, specifically when performed globally and subject to confidentiality in some parts of the world, is highly expensive and the huge size of data on countries around the globe demanding the dedication of well-trained big teams to collect and organize these data is a time consuming process. There are also several issues related to the data obtained from the World Bank's database that make the adoption and utilization very complicated and challenging.

One major issue is that not all needed location decision factors could be directly found in the World Bank's database or other global indices. In this case, the unavailable factors are represented by one or more sub-indices and the numerical data of these sub-indices are collected and combined with the numerical data collection of remaining factors.

Another issue of numerical data in the World Bank's database is the missing data of some or all factors for some countries. Ideally, this issue could be resolved as follows:

- Data are missing for all the time periods; then the associated country(s) is/are excluded, because such countries are most likely either to have no significant data to share or they lack political transparency.
- Data are missing for several time periods; in this case, capturing the missing data could be done by first looking up the data in other global database sources. The missing data also could be forecasted based on available data of previous time periods.

Moreover, the numerical data of the decision-making factors exist in the World Bank's database with different ranges of values; some of them are wider than others. Therefore, it is important for these data to be refined before they can be used in the clustering approach to form the desired groups of countries. To insure data integrity and in order to prevent getting conditioned by features with a wider range of possible values when computing coefficients, the numerical data need first to be normalized. In this research, the approach used to normalize data is the feature scaling (min-max scaling) that is typically calculated using the formula

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

The resulted normalized data through this approach are scaled to a fixed range between (0-1) with a smaller standard deviation to help suppress the effect of outliers.

4.2 Weight assigning to location decision factors

All determined factors are critical and important for entrepreneurs to choose a best-fit location for their small or medium starting-up ventures. However, scholars have stressed some of these factors more than others. Therefore, the decision-making process could be improved by making these criteria more explicit. Assigning a weight to each identified factor can be based on how strongly entrepreneurship scholars emphasized it in their research, i.e., the more

entrepreneurship scholars emphasized a location's decision-making factors, the higher weight it is given. Assigning the weights to these factors is a good way to find mismatches on expectations. It also helps decision makers to be less subjective and be more objective in evaluating available alternatives.

Taking into account the literature discussed in Chapter Two and the decision-making factors identified in Chapter Three, weights can be potentially assigned to the defined location decision factors as follows:

Table (4.1) Weights assigned to entrepreneurial facility location decision factors

#	Decision-making factors (attributes)	Weight (%)
1	Cost of business start-up procedures	1.56
2	Start-up procedures to register a business	0.9
3	Time required to start a business	0.8
4	Patent applications	0.7
5	Trademark applications	0.6
6	Charges for the use of intellectual property	0.5
7	Firms using banks to finance investment	5
8	Lending interest rate	8
9	Foreign direct investment	6
10	Total tax rate	4
11	Profit tax	3
12	Taxes on goods and services	2
13	Exports of goods and services	0.4
14	Trade in services	0.3
15	Net official development assistance and official aid received	0.2
16	Labor force with tertiary education	0.09
17	Secondary education, vocational pupils	0.08
18	Government expenditure on education	0.1
19	Wage and salaried workers	0.07
20	Unemployment	0.06
21	Investment in energy	0.05
22	Investment in telecoms	0.05
23	Investment in transport	0.05
24	Investment in water and sanitation	0.05

25	High-technology exports	10.4
26	Internet users	9
27	Fixed broadband subscriptions	9
28	Research and development expenditure	14
39	Researchers in R&D	11
30	Manufacturing, value added	12
31	Listed domestic companies	0.04

An average rank is then applied on the weighted numerical data and subsequently, the top one hundred countries in the resulting ranked list of countries that will be also compared with the lists identified through credible entrepreneurship indices, e.g., the Global Entrepreneurship Index of the Global Entrepreneurship and Development Institute (GEDI), is adopted as preferable locations for entrepreneurs to establish their start-up facilities.

4.3 Data collection challenges and implications

In spite of applying all of the preceding steps in order to refine the collected data and prepare them to be implemented as inputs for the research methodology, the problem of the unavailability of significant and critical data for some countries inhibits the correct interpretation of the entrepreneurial attraction factors' impacts on the location decision-process of entrepreneurs. Thus, to illustrate the methodology of this research in full, a hypothetical case study is discussed in the following sections.

Furthermore, in order to add more sense to the generated results, a real-time demonstration of the clustering approach will be conducted, taking into account installing only the data for available decision-making factors that are complete and with no missing values.

4.4 Model development

The proposed model in this research is based on a hierarchical clustering algorithm that starts by singular objects. Then it gradually gathers them into homogenous groups according to

their similarities in regard to location-attractiveness factors until eventually one large cluster of objects can be formed at the last iteration of the algorithm. Moreover, the developed clustering technique has to be stopped before merging all generated clusters for determining the required number of clusters instead of one unique cluster.

As discussed earlier in section 2.4.1, grouping the considered dataset may be done in various fashions in accordance to the selected clustering method. In fact, it is the type of desired output that actually dictates the selection of a particular method. Furthermore, there are also several unique characteristics that most likely affect the selection of the clustering method, including the performance of the method with specific data type, the available hardware and software facilities for the selected method, and the size of the dataset the method can handle.

Depending on most important categories of dataset grouping. Table (4.2) presents a basic comparison between existing multi-criteria decision-making and the proposed approaches.

Clustering algorithms mostly consist of three main components:

- Objects
- Attributes
- Similarity coefficient

Similarly, components of the clustering model in this research are the objects, the attributes, and the similarity coefficient.

Objects: of the proposed model are the countries to be processed by the clustering technique in order to be combined together and form homogenous groups. Like other clustering algorithms, the objects (countries) in the model introduced in this research are grouped together

such that the objects in one group are similar to each other whereas they differ from other objects belonging to other groups.

Attributes: are the set of variables upon which available attributes are compared and the similarities among them are also measured to choose the best alternatives. Attributes are the backbone of clustering techniques and the specified set should be comprehensive and contain all critical variables. Failing to include important attributes will most likely result in the formation of clusters that are inefficient or nonhomogeneous, i.e., assigning similar objects into separate groups. Attributes of the proposed model are the location decision-making factors that have been identified in the previous chapter, in order to create a comprehensive list that considers all aspects of the desired decision.

Similarity Coefficient: is generally the mathematical function by which the similarities of two or more objects are measured based on the values of attributes. There are many similarity coefficients suggested by researchers; however, choosing the similarity coefficient type depends on the characteristics of attributes in comparison, as well as the desired clustering of objects mentioned.

Furthermore, the notation that can be used in the formation and development of the introduced model is given by the following:

i and j are any two countries to be compared as potential locations

a_{ij} is any attribute used for the comparison between country i and country j is conducted

m number of the countries to be listed as alternatives (rows of the similarity matrix)

$\{a_{ij}\}$ country-attribute incidence matrix =

$$\begin{cases} 1 & if\ the\ ath\ attribute\ exists\ within\ the\ same\ range\ of\ strngth\ in\ both\ i\ and\ j\ countries \\ 0 & otherwise \end{cases}$$

n total number of attributes exist in countries = $\sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij}$

S_{ij} Jaccard similarity coefficient between country i and country j

$d(i,j)$ Euclidean or CityBlock distance between country i and country j

As implied to in Chapter Two, the proposed clustering analysis model is a derivation of a proven clustering approach utilized in the field of manufacturing to study the formation of clusters of machine cells visited by part families based on specified attributes of the parts. This approach was modified so that a similarity coefficient-based clustering algorithm, namely the complete linkage-clustering method (CLINK), is implemented to create clusters of similar countries that have the potential to offer the best locations to start up entrepreneurial ventures with the consideration of factors that are appealing to entrepreneurs. However, other coefficient-based clustering algorithms will be also applied to obtain a solid base to review the differences between those approaches (if any). A comparison of the clustering approach that is used to form the clusters of machine cells and the proposed model to create clusters of similar countries in their attractiveness to entrepreneurial firms is shown below.

Table (4.2) Comparison of components of the clustering approach in manufacturing and the components of the proposed clustering model

Component	Clustering approach in manufacturing	Proposed clustering model in entrepreneurship
Objects	Machine cells visited by part families	Countries that have the potential to accommodate entrepreneurial firms
Attributes	Characteristics of parts upon which they are classified into families	Factors that attract entrepreneurship to the locations
Similarity coefficient	JSC, Euclidean distance, etc.	JSC, Euclidean distance, CityBlock
Expected outcomes	Clusters of machine cells visited by part families based on specified attributes of the parts	Clusters of countries that have similar attractiveness for entrepreneurs based on their location tempting attributes

4.4.1 Transforming the numerical values of attributes into the proper formats

The following section illustrates the procedure through which the similarity coefficients are obtained using the Jaccard, Euclidean and the CityBlock models.

a) Jaccard Similarity Coefficient (JSC)

In Chapter Two of this research, the JSC approach for the machine-clustering example was calculated depending on the number of parts visiting each machine. However, the JSC of the developed model is based on the range of the strength level of each of the decision-making factors. Here also the attributes are set to be binary and yielded possibilities for each pair of countries as 1-1, 1-0, 0-1, and 0-0.

Since the JSC is solely designed for binary variables (i.e., take 0 or 1 values), therefore, all attributes have to be transformed into binary variables. There are only four possible outcomes resulting from the JSC calculation for any two countries: 0-0, 0-1, 1-0, and 1-1. The first digit in this notation represents the binary value (likelihood) of the attribute for the first country, while the second digit in this notation represents the binary value (likelihood) of the attribute for the second country. The similarity between any two countries (objects) increases with an increase in the JSC value, which can only be between 0 and 1. The maximum value is obtained when the two considered countries (objects) have completely identical values for each attribute, and yields a minimum value (zero) when the countries have dissimilar values for each of the considered attributes.

In this research, the following steps are carried out to transform the attribute values into binary numbers to enable the determination of the JSCs.

- The minimum and maximum values of each attributes are determined
- The ranges are calculated for each attribute (range = maximum value – minimum value)

- Each range is divided into four equal intervals (interval length = range / 4)
- Four corresponding binary dummy variables are created with respect to the above interval. The value of each attribute for each country is assessed with respect to the above set intervals, and a dummy variable is assigned as follows: the attribute value = 1 only for the interval in which the actual attribute value lies, and otherwise the value is set to zero.
- The expected possibilities for the decision-making factors in the Jaccard Similarity Coefficient (JSC) to be applied into the introduced model are shown by the contingency matrix in Table (4.3) where the values of a, b, c, and d are binary number i.e a = 1, b = 1 or 0, c = 0 or 1, and d = 0.

Table (4.3) The assignments of each attribute for any two countries i and j.

		Country *j*	
		1	0
Country *i*	1	$a = 1$	$b = 1$ or 0
	0	$c = 1$ or 0	$d = 0$

Where a is 1 when a given factor (attribute) belongs to the interval of high level of strength in both countries *i* and *j*, b is 1 or 0 when a given factor (attribute) does (or does not) belong to the interval of high level of strength in either country *i or j*, c is 1 or 0 when a given factor (attribute) does (or does not) belong to the interval of high level of strength in either country *i or j*, and d is 0 when a given factor (attribute) does not belong to the interval of high level of strength in either country *i* or country *j*.

From the contingency matrix shown in Table (4.3), JSC is calculated by the following formula:

$$s_{ij} = \frac{a}{a+b+c}, \quad 0 \leqslant s_{ij} \leqslant 1 \qquad \text{(Yin and Yasuda, 2006)}$$

Furthermore, as it was presented in the Section 2.4.2, according to McAuley the Jaccard similarity coefficient between two machines in cellular manufacturing is defined as the number of parts visiting both machines divided by the number of parts that visit either of the two machines and therefore it is calculated by the following mathematical formula:

$$s_{ij} = \frac{\sum_{k=1}^{N} X_{ijk}}{\sum_{k=1}^{N} Y_{ijk}},$$

where:

S_{ij} = the similarity coefficient between machines *i* and *j*

$$X_{ijk} = \begin{cases} 1 & if\ both\ machines\ i\ and\ j\ are\ visited\ by\ part\ type\ k \\ 0 & otherwise \end{cases}$$

$$Y_{ijk} = \begin{cases} 1 & if\ at\ least\ one\ of\ the\ machines\ i\ or\ j\ is\ visited\ by\ part\ type\ k \\ 0 & otherwise \end{cases}$$

and N = number of parts

Similarly, in the global facility location for the entrepreneurial firms, the Jaccard similarity coefficient between any two countries (i and j), is defined as the number of attributes that are strong in both countries divided by the number of strong attributes that are strong in either of the two countries as shown by the following general formula:

$$s_{ij} = \frac{\sum_{k=1}^{N} X_{ijk}}{\sum_{k=1}^{N} Y_{ijk}}, \text{ where}$$

S_{ij} = the similarity coefficient between countries *i* and *j*

$$X_{ijk} = \begin{cases} 1 & if\ both\ countries\ i\ and\ j\ have\ a\ strong\ level\ of\ attribute\ k \\ 0 & otherwise \end{cases}$$

$$Y_{ijk} = \begin{cases} 1 & if\ at\ least\ one\ of\ the\ countries\ i\ or\ j\ has\ a\ strong\ level\ of\ attribute\ k \\ 0 & otherwise \end{cases}$$

and N = number of attributes

b) Euclidean distance

The Euclidean distance is used to measure the similarity between two clusters. For two countries i and j where $i \epsilon$ cluster t and $j \epsilon$ cluster v, k is any attribute and N is the number of attributes in the model, the Euclidean distance is calculated by

$$d(t,v) = \left(\sum_{k=1}^{N} (x_{ik} - x_{jk})^2 \right)^{1/2}$$

For proper implementation of the Euclidean distance in the developed model, following steps are conducted:

- Collecting the numerical data of the decision-making attributes from the related World Bank's database
- Performing the feature normalizing process using the min-max scaling approach to prevent the influence by features with a wider range of possible values when computing coefficients
- Obtaining the resulting normalized attributes (between 0-1) with a smaller standard deviation to help suppress the effect of outliers
- Calculating the Euclidean distance using the above given formula, where the two starting clusters (t and v) are any two randomly selected countries.

c) City Block distance

The calculation of the CityBlock distance requires similar steps as those for the Euclidean distance prior using it in the model. These steps are:

- Collecting the numerical data of the decision-making attributes from the related World Bank's database
- Performing the feature normalizing process using the min-max scaling approach to prevent the influence by features with a wider range of possible values when computing coefficients
- Obtaining the resulting normalized attributes (between 0-1) with a smaller standard deviation to help suppress the effect of outliers
- Calculating the CityBlock between two countries $i \in$ cluster t and $j \in$ cluster v using the formula

$$d(t, v) = \sum_{k=1}^{N} \left| x_{ik} - x_{jk} \right|,$$

where k = (1, 2, … , *N*) is an attribute, i and j are the countries.

4.4.2 Clustering technique

After computing the similarity coefficients through the JSC method, or the distances obtained through the Euclidean or the CityBlock approaches, the candidate countries to accommodate the location of new small or medium businesses are classified using the complete linkage clustering (CLINK) technique. CLINK, which is a similarity coefficient-based clustering methodology, was applied in this research due to the following reasons (Gupta and Seifoddini, 1990):

- The least similar pair between two clusters is used to determine the inter-cluster similarity
- Resulting clusters are small and tightly bound
- It prevents the merging of two clusters together for only a high level of similarity between two members while other members are dissimilar
- It is computer software-friendly (e.g., MATLAB has inbuilt CLINK functions)

CLINK algorithm demands adopting a clustering method to obtain clusters in which the addition of an entity to a cluster must not require that the entity is highly similar to any member of that cluster, i.e., preventing the chaining reaction (formation of clusters that can tend to resemble long chains).

CLINK starts with computing the similarity coefficient for each pair of the object groups, where as a starting point; each individual object (country) is initially considered to be its own cluster and the standard steps for executing The CLINK algorithm are:

1. Set up the similarity matrix by calculating the similarity coefficient for each pair of groups (countries).
2. Determine the groups of countries with the maximum similarity coefficient and put them together.
3. Eliminate the rows corresponding with the country groups that were grouped together.
4. Add a new row to the matrix for the resulting new country group and compute the similarity coefficient using the formula $S_{tv} = Min\{S_{ij}\}\, i \in t \,\&\, j \in v$, where t is the new cluster of countries and v is the other clusters, i.e. the countries that lie in the various clusters are grouped together base on the minimum existing similarity between those countries.

5. Repeat the steps from step 2 to step 4.
6. The algorithm terminates when the number of machine groups that was previously determined is achieved.

It is also required to apply the obtained JSC, Euclidean and CityBlock coefficients (or distances) to the CLINK methodology. This can be carried out using the following equations:

- For the Jaccard Similarity Coefficient (JSC)

$$d(t,v) = Min\{S_{ij}\}\, i \in t \;\&\; j \in v),$$

- For the Euclidean distance

$$d(t,v) = Min\left\{\left(\sum_{k=1}^{N}(x_{ik} - x_{jk})^2\right)^{1/2}\right\}$$

- For the CityBlock distance

$$d(t,v) = Min\left\{\sum_{k=1}^{N}|x_{ik} - x_{jk}|\right\}$$

According to Anandan and as it was illustrated in Section 2.4.2, the distance between two clusters X and Y is computed as the maximum distance between any two points $x \in X$ and $y \in Y$ in the two clusters, i.e., $d(X,Y) = \max_{x \in X, y \in Y} d(x,y)$

Similarly following the calculations of the similarity coefficients to be applied in the developed model, the CLINK algorithm hierarchically forms the clusters (dendrogram) by considering the maximum distance between any two countries $i \in t$ and $j \in v$ in the two clusters t and v as shown in the following equation:

$$D(t,v) = \max_{i \in t, j \in v} d(i,j)$$

4.4.3 Obtaining the related dendrograms and clusters through MATLAB

The next step after the identification of the entrepreneurial location decision factors, collecting required numerical values and setting up the data to study similarities and

dissimilarities of the considered locations is performing a clustering analysis using the MATLAB. The functions included within the built-in codes in MATLAB support agglomerative clustering and perform all of the required steps.

Following are the important steps of MATLAB to perform the hierarchical clustering in order to determine the distinct groups of countries based on their similarities and dissimilarities in attracting the entrepreneurial facilities/activities.

- Finding the similarity and dissimilarity between every pair of objects in the considered set of data.

  The (pdist) function calculates the distance between objects (countries). Unless otherwise specified differently, the (pdist) function is set to calculate the Euclidean distance between objects (countries) by default.

- Grouping the countries into a binary, hierarchical cluster tree.

  Using the generated information about distance in the last step to determine the proximity of objects (countries) to each other, the (linkage) function links pairs of objects (countries) that are in close proximity. These newly formed clusters are then grouped into larger clusters leading to the formation of a hierarchical tree.

- Determining the number of clusters by detecting natural groupings in the hierarchical tree or by cutting it off at an arbitrary point.

  The function (dendrogram (tree, P)) is used to obtain the desired number of clusters and related tree diagram (dendrogram).

4.5 Validating the developed model

In this research, the developed model of clustering analysis to help promote the location decision process among entrepreneurs to select the best-fit site for their starting-up ventures is

displayed through two main approaches. The first approach is a hypothetical case study in which the implemented similarity coefficient is the Jaccard similarity coefficient and the adopted clustering analysis technique is the Complete Linkage Clustering (CLINK) algorithm. The second approach is applying the Complete Linkage Clustering (CLINK) algorithm upon a real-world sample with available complete numerical data obtained from the World Bank's database, but with the Euclidean distance as the similarity coefficient to be installed.

Moreover, several similarity coefficients with more than one clustering algorithm will be applied for the clustering analysis of a large size real-world sample in order to further examine the validity of the proposed model.

CHAPTER FIVE

Applications of the Developed Model and Methodology

In this chapter, the developed model of clustering analysis to help promote the location decision process among entrepreneurs to select the best-fit site for their starting-up ventures is tested through various distinctive approaches. The first is a hypothetical case study in which the similarity coefficient to be used is the Jaccard similarity coefficient that can be plugged into the Complete Linkage Clustering (CLINK) algorithm to create the clusters of countries according to their appeal to entrepreneurial small and medium starting-up enterprises.

The second approach is applying the developed model onto a real-world sample that has complete numerical data obtained from the World Bank's database. The cluster analysis technique is also the Complete Linkage Clustering (CLINK) algorithm, except that the similarity coefficient is computed using the Euclidean distance.

The last approach that will be used to test both the validity and flexibility of the model is carried out first through the application of the Complete Linkage Clustering (CLINK) algorithm with Euclidean distance as the similarity coefficient for the clustering analysis of a large size real-world sample, and second by applying several similarity coefficients with different similarity coefficient-based clustering algorithms to the clustering analysis of the same large size real-world sample.

5.1 Hypothetical Case Study

Assuming that after identifying a competitive business idea that can be interpreted into highly desired products or distinctive services to be provided to interested customers at a global level, the entrepreneur develops a good business plan in which she/he forecasts a prospective

market internationally for their products or services and intends to take advantage of the evolving global markets phenomenon, the rapid advancement in communication, and the improvements and easiness of cross-borders transportation. In the business plan, the entrepreneur considers all possible success factors in the international markets via embedding a global vision that reflects a deep understanding of the complexity of international markets. Furthermore, the entrepreneur also prepares comprehensive business propositions that take into account the formation of strong production and distribution networks, strengthening cross-cultural competence, and aligning physical and human resources to facilitate the entry of their business into the selected foreign market.

Subsequently, in order to improve the selection process of which foreign market to establish the new venture within and to achieve a more efficient location decision, the involved entrepreneur(s) has to conduct an extensive study to create a list of candidate countries that are most likely to accommodate the new-born facility. Then, the entrepreneur can conduct a comparison process between these alternate countries based on a set of attributes to select the best-fit location among the specified countries.

For the hypothetical case study it is assumed that the entrepreneur(s) would implement the clustering analysis model developed in the research to decide upon the optimal location to start up the business in an ideal case where all needed numerical data is available. Both the name of countries and the provided data are hypothesized throughout the case study.

5.1.1 Developing the list of candidate countries

This stage begins by collecting as much comprehensive information as possible about different countries that might be suitable to host the entrepreneurial facility. Then, an initial

analysis of these candidate countries is conducted based on the needs and available resources of the business to finalize the list of 20 countries that are most likely to include the requirements to establish the firm within. An illustration of the listed 20 candidate countries (hypothetically named Country (1), Country (2), ……, and Country (20)) is given in Table (5.1).

Table (5.1) Final list of the hypothetical candidate countries

#	Country
1	Country (1)
2	Country (2)
3	Country (3)
4	Country (4)
5	Country (5)
6	Country (6)
7	Country (7)
8	Country (8)
9	Country (9)
10	Country (10)
11	Country (11)
12	Country (12)
13	Country (13)
14	Country (14)
15	Country (15)
16	Country (16)
17	Country (17)
18	Country (18)

19	Country (19)
20	Country (20)

5.1.2 Identifying the decision making factors

The second stage in applying the model of the research is to associate the listed countries in stage one according to their similarities and dissimilarities in order to ultimately form clusters of countries to provide the entrepreneur with a highly efficient means to select the best foreign market to establish her/his new business. Assigning similar countries into clusters requires specifying criteria to measure how similar or dissimilar the investigated countries are. The measuring criteria must be carefully selected to cover the various aspects of the entrepreneurial facilities' location decision problem. In this research, as well as in the hypothetical case study, these criteria are the attributes or the decision-making factors that have been defined in Chapter Three of this research. A summary of the most critical decision-making factors is given in Table (5.2).

Table (5.2) List of location decision-making factors and associated sub-factors

Main factor	Decision-making sub-factors (attributes)
Business start-up cost and procedure	Cost of business start-up procedures
	Start-up procedures to register a business
	Time required to start a business
	Patent applications
	Trademark applications
	Charges for the use of intellectual property
Financing small and medium enterprises	Firms using banks to finance investment
	Lending interest rate
	Foreign direct investment

Tax rates and structure	Total tax rate Profit tax Taxes on goods and services
Governmental regulations and policies	Exports of goods and services Trade in services Net official development assistance and official aid received
Labor and skills	Labor force with tertiary education Secondary education, vocational pupils Government expenditure on education Wage and salaried workers Unemployment
Infrastructure	Investment in energy Investment in telecoms Investment in transport Investment in water and sanitation
Technology advancement	High-technology exports Internet users Fixed broadband subscriptions Research and development expenditure Researchers in research and development Manufacturing, value added
Competition	Listed domestic companies

5.1.3 Processing of data collection and setup

The next stage following development of the list of suitable countries to accommodate the entrepreneurial facility and specifying the most important location decision-making factors to measure similarities of these countries is collecting data that provide numerical values upon the decision-making factors for each one of the countries in comparison. Data collection can be conducted through various data collection techniques, such as dispatching well-prepared surveys,

revising authentic databases, interviewing experienced entrepreneurs, and exploring previously conducted studies and literature. However, some of the needed data might be unavailable or hard to obtain directly using any procedure of data collection. Then, this type of data is represented by one or more related sub-indices and the numerical data of these sub-factors are considered in the model.

As for the considered hypothetical case study, the required numerical values are generated randomly based on the ranges of numerical data that are found in the World Bank's indicators directory and they are assumed to represent the data collected upon location decision-making factors. The randomly created data for each of the decision-making factors are shown independently in the following tables.

Table (5.3) Hypothetical data for the sub-factors of business start-up cost and procedure

Country	Cost of business start-up procedures	Start-up procedures to register a business	Time required to start a business	Patent applications	Trademark applications	Charges for the use of intellectual property
Country (1)	2.2	4	5.5	193	2261	1492290707
Country (2)	3.1	8	22	224	3415	1879594723
Country (3)	1.9	1	5	30174	28370	10229810323
Country (4)	0.2	6	10	8579	8146	3792969133
Country (5)	0.8	2	2	26656	3739	52812610
Country (6)	1.2	3	2.5	167	22878	4051970239
Country (7)	3.6	7	9	16149	7570	202200000

Country (8)	6.2	5	11	13690	2580	2710876.798
Country (9)	6.6	12	31	631	17894	12351406708
Country (10)	0.7	5	13	4984	1653	445830662.3
Country (11)	0.4	3	5	258	5033	110796021.6
Country (12)	7.2	2	2	25925	4182	19348328.75
Country (13)	0.3	4	4	32362	2120	45136670.91
Country (14)	8.3	5	5	3065	4207	22040000
Country (15)	0.6	8	4	14234	7182	45785716607
Country (16)	0.5	5	10	6787	13454	3046393.956
Country (17)	4.6	4	8	357	18216	3971506597
Country (18)	9.2	3	5.5	44983	17520	500583339.7
Country (19)	5.0	10	9	124	2477	1728192135
Country (20)	3.4	6	8	869	4492	241500000

Table (5.4) Hypothetical data for the sub-factors of financing small and medium enterprises

Country	**Firms using banks to finance investment**	**Lending interest rate**	**Foreign direct investment**
Country (1)	17.9	5.80	9079291878
Country (2)	30.3	12.8	4303046353
Country (3)	38.4	3.00	-789862234.7
Country (4)	29.9	5.95	-370016674.1

Country (5)	2.0	9.12	-220712949.2
Country (6)	12.8	7.45	-388308435.2
Country (7)	6.7	8.09	-164301885.3
Country (8)	6.4	8.27	-299000000
Country (9)	19.2	5.60	-755578832
Country (10)	2.8	16.41	5707967033
Country (11)	8.0	11.14	-363269189
Country (12)	28.3	4.76	-520988091.8
Country (13)	5.4	4.27	-251200000
Country (14)	11.6	3.25	1446349192
Country (15)	22.1	7.74	-72197486.07
Country (16)	7.4	15.50	-566248722.3
Country (17)	20.3	11.91	789477357.4
Country (18)	9.2	3.91	-19782558.05
Country (19)	35.0	17.22	-877423257.6
Country (20)	13.4	6.77	-260785960

Table (5.5) Hypothetical data for the sub-factors of tax rates and structure

Country	**Total tax rate**	**Profit tax**	**Taxes on goods and services**
Country (1)	26	20.3	16.68

Country (2)	52	15.4	9.15
Country (3)	21	3.9	14.09
Country (4)	28.6	15.9	15.24
Country (5)	7.4	5.5	11.03
Country (6)	47.3	26.1	0.14
Country (7)	13.5	11.9	10.61
Country (8)	12.8	7.4	6.43
Country (9)	49.9	9.1	17.98
Country (10)	39.8	8.4	9.10
Country (11)	16.4	9.5	22.56
Country (12)	35	10.7	0.37
Country (13)	39.7	15.1	18.70
Country (14)	31.5	12.9	6.55
Country (15)	26.9	14.3	6.17
Country (16)	39.2	4.9	4.60
Country (17)	48	9.3	15.14
Country (18)	45	21.1	10.76
Country (19)	15.8	14.4	2.75
Country (20)	13.6	19.9	13.31

Table (5.6) Hypothetical data for the sub-factors of governmental regulations and policies

Country	Exports of goods and services	Trade in services	Net official development assistance and official aid received
Country (1)	53.68	39.98	266670000
Country (2)	53.62	27.55	89720000
Country (3)	31.56	10.82	243530000
Country (4)	47.86	9.68	77200000
Country (5)	20.91	25.78	101780000
Country (6)	57.20	8.10	107020000
Country (7)	19.91	17.65	285820000
Country (8)	39.63	21.77	515690000
Country (9)	32.95	39.69	461910000
Country (10)	43.28	23.87	773570000
Country (11)	42.88	19.52	122810000
Country (12)	58.02	28.59	141800000
Country (13)	42.11	24.69	533310000
Country (14)	83.14	26.84	127750000
Country (15)	75.05	36.22	39040000
Country (16)	79.64	33.61	296050000
Country (17)	15.19	29.02	646120000

Country (18)	12.35	30.94	675090000
Country (19)	22.26	13.74	138370000
Country (20)	50.99	14.71	364560000

Table (5.7) Hypothetical data for the sub-factors of labor and skills

Country	**Labor force with tertiary education**	**Secondary education, vocational pupils**	**Government expenditure on education**	**Wage and salaried workers**	**Unemployment**
Country (1)	31.6	134687	19.15	91.0	7.0
Country (2)	20.0	259553.07	19.95	86.6	4.9
Country (3)	34.0	359191	12.30	69.4	13.0
Country (4)	21.5	164240	20.30	71.9	5.2
Country (5)	29.4	245912	14.77	85.1	28.9
Country (6)	37.1	817938	21.67	88.19	2.7
Country (7)	26.6	335482	15.61	87.9	11.7
Country (8)	33.7	164584	20.57	80.9	12.9
Country (9)	37.4	6970	16.86	90.9	9.8
Country (10)	38.2	89993	9.58	84.7	7.9
Country (11)	21.0	7204	18.45	67.8	8.6
Country (12)	37.2	252191	15.20	63.0	3.3
Country (13)	19.8	29562	11.55	78.1	4.4

Country (14)	25.5	271345	17.24	33.5	5.8
Country (15)	34.3	155326	13.99	81.9	7.3
Country (16)	41.2	49208	14.98	79.4	27.3
Country (17)	29.1	21350	18.46	82.1	4.9
Country (18)	19.8	28440	20.64	68.8	14.5
Country (19)	16.3	177014	19.03	34.8	17.2
Country (20)	31.1	874889	13.82	83.9	11.9

Table (5.8) Hypothetical data for the sub-factors of infrastructure

Country	**Investment in energy**	**Investment in telecoms**	**Investment in transport**	**Investment in water and sanitation**
Country (1)	3440700000	1452900000	677600000	116760000
Country (2)	780450000	560800000	79000000	2947100000
Country (3)	461100000	74100000	4475760000	129000000
Country (4)	1087500000	306100000	101000000	135100000
Country (5)	40000000	85096000	63000000	86150000
Country (6)	1790000000	227200000	3100000	140000000
Country (7)	2022000000	187740000	11768100000	3162250000
Country (8)	9651800000	7969600000	275500000	62250000
Country (9)	78300000	17800000	381000000	140065700
Country (10)	3537710000	184800000	788600000	156500000

Country (11)	774600000	667300000	916000000	160410000
Country (12)	2600000000	60400000	30500000	2247100000
Country (13)	1148300000	354000000	4587300000	2447100000
Country (14)	177000000	344200000	3215000000	97150000
Country (15)	125000000	130000000	73400000	55160000
Country (16)	2873400000	10209900000	80000000	93400000
Country (17)	518500000	43400000	25000000	66150000
Country (18)	3392800000	1706000000	4040500000	1455100000
Country (19)	269000000	99500000	4400000000	2548100000
Country (20)	421200000	924800000	1149300000	1345700000

Table (5.9) Hypothetical data for the sub-factors of technology advancement

Country	High-technology exports	Internet users	Fixed broadband subscriptions	Research and development expenditure	Researc-hers in R&D	Manufactur-ing, value added
Country (1)	9185071604	95.99	41.38	2.98	6730.39	41104921187
Country (2)	18412394058	81.00	27.54	2.84	4564.94	72591175250
Country (3)	29136849244	87.12	34.19	1.73	5181.19	41678660610
Country (4)	3074242429	54.89	12.93	3.40	6437.73	22152281049
Country (5)	117522964	68.06	16.19	2.09	1551.97	1203218179
Country (6)	4565211317	84.56	25.76	0.64	7482.34	92768040423

Country (7)	766682600	90.99	21.39	2.68	3096.11	17630261434
Country (8)	3724745015	92.38	32.30	3.55	2804.18	830701871
Country (9)	22066017	78.70	45.97	0.82	6193.87	217987528.2
Country (10)	84074873	63.21	28.36	0.69	828.14	17727816247
Country (11)	854961530	61.00	19.83	1.49	2168.34	3590080326
Country (12)	71872403	48.90	12.15	3.93	1552.67	1908836609
Country (13)	21682663	75.83	24.74	0.66	1894.75	96953721588
Country (14)	925175276	46.60	14.71	1.30	2719.07	3368709627
Country (15)	14300836	93.17	41.02	2.16	3505.96	933185613.7
Country (16)	4346223	49.28	11.51	0.99	3111.49	131419225
Country (17)	33901233425	34.89	8.21	2.32	2389.39	78303996986
Country (18)	60371906718	67.50	10.14	1.29	4138.88	45051624471
Country (19)	14470677145	76.13	27.35	2.92	1853.83	14899884610
Country (20)	70412348	31.70	3.68	0.84	1820.22	32919454347

Table (5.10) Hypothetical data for the sub-factor of competition

Country	**Listed domestic companies**
Country (1)	174
Country (2)	70
Country (3)	3876

Country (4)	74
Country (5)	32
Country (6)	1959
Country (7)	329
Country (8)	189
Country (9)	267
Country (10)	284
Country (11)	133
Country (12)	31
Country (13)	3167
Country (14)	105
Country (15)	36
Country (16)	502
Country (17)	921
Country (18)	51
Country (19)	234
Country (20)	573

5.1.4 Conversion of data's real values into binary variables

This stage involves transforming the real values of numerical data of the decision-making factors into the binary variable form (with values of only 0 or 1) prior to installing them into the

JSC calculations in order to be used to measure similarities of attributes in the clustering analysis model of the hypothetical case study. To convert the data into binary variables form, typical steps that have been explained in section 4.5.1 are applied on the pre-defined location decision-making factors.

5.1.4.1 Data conversion into binary variables for the *cost of business start-up procedures*

a. Determining minimum and maximum values of the attributed cost of business start-up procedures among the different values within the different alternatives;

   min= 0.2 and max= 9.2

b. Calculating the value range (range = maximum value – minimum value);

   range= 9.2 – 0.2 = 9

c. Dividing the range into four equal intervals (length = range / 4);

   length= 9/4 = 2.25 and the yielded intervals are:

   - Interval 1: [0.2, 2.45)
   - Interval 2: [2.45, 4.7)
   - Interval 3: [4.7, 6.95)
   - Interval 4: [6.95, 9.2]

d. Creating four corresponding new binary variables to the intervals; X_1, X_2, X_3, *and* X_4. The binary variable that is correspondent to the interval contains the value of cost of business start-up procedures is assigned the value 1, while others are assigned the value 0. The corresponding binary variables X_1, X_2, X_3, *and* X_4 are defined as follows:

   - $X_1 = \begin{cases} 1 \; if \; the \; value \; falls \; within \; the \; interval \; [0.2 \, , 2.45) \\ 0 \quad otherwise \end{cases}$
   - $X_2 = \begin{cases} 1 \; if \; the \; value \; falls \; within \; the \; interval \; [2.45 \, , 4.7) \\ 0 \quad otherwise \end{cases}$

- $X_3 = \begin{cases} 1 \; if \; the \; value \; falls \; within \; the \; interval \; [4.7 \, , 6.95) \\ 0 \quad otherwise \end{cases}$
- $X_4 = \begin{cases} 1 \; if \; the \; value \; falls \; within \; the \; interval \; [6.95 \, , 9.2] \\ 0 \quad otherwise \end{cases}$

The resulted binary variables form the location decision-making factor; the cost of business start-up procedures is shown in Table (5.11).

Table (5.11) Binary variables of the sub-factor: cost of business start-up procedures

Country	**Cost of start-up procedures**	X_1	X_2	X_3	X_4
		$[0.2 \, , 2.45)$	$[2.45 \, , 4.7)$	$[4.7 \, , 6.95)$	$[6.95 \, , 9.2]$
Country (1)	2.2	1	0	0	0
Country (2)	3.1	0	1	0	0
Country (3)	1.9	1	0	0	0
Country (4)	0.2	1	0	0	0
Country (5)	0.8	1	0	0	0
Country (6)	1.2	1	0	0	0
Country (7)	3.6	0	1	0	0
Country (8)	6.2	0	0	1	0
Country (9)	6.6	0	0	1	0
Country (10)	0.7	1	0	0	0

Country (11)	0.4	1	0	0	0
Country (12)	7.2	0	0	0	1
Country (13)	0.3	1	0	0	0
Country (14)	8.3	0	0	0	1
Country (15)	0.6	1	0	0	0
Country (16)	0.5	1	0	0	0
Country (17)	4.6	0	1	0	0
Country (18)	9.2	0	0	0	1
Country (19)	5	0	0	1	0
Country (20)	3.4	0	1	0	0

5.1.4.2 Data conversion into binary variables form for the remaining factors

Similarly, data of the remaining location decision-making factors are transformed into the binary variables form through applying the same procedure.

5.1.5 Implementing the clustering analysis model

At this stage all the data must have been converted into binary variables. Therefore, the set up of the required data is completed and becomes ready to be installed in the developed clustering analysis model in which the complete linkage clustering method (CLINK) is adopted.

By implementing the developed clustering model, the considered countries are grouped into six distinctive clusters. Each cluster combines the countries that are most similar with respect to the specified critical location decision-making factors. The resulted dendrogram from the implementation of the developed clustering analysis model is shown in Figure (5.1).

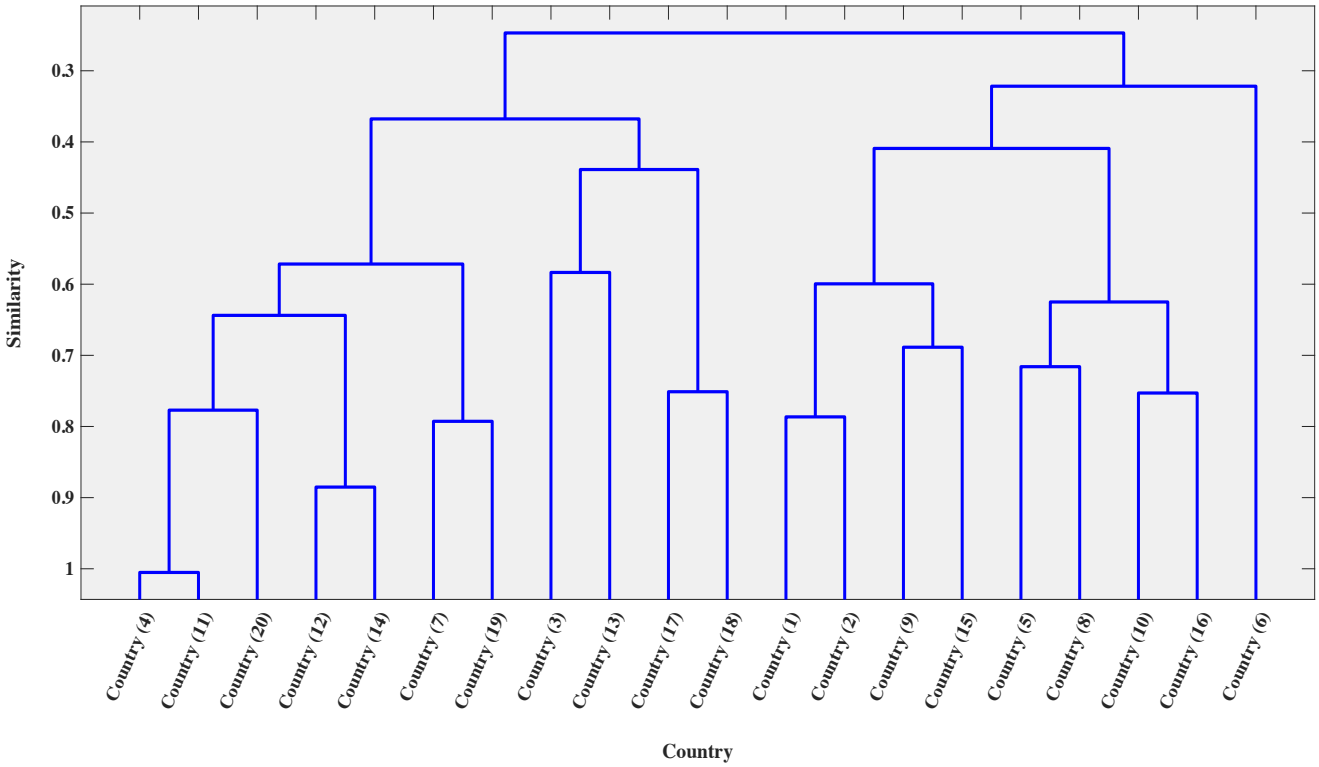

Figure (5.1) Dendrogram of the developed model for the hypothetical case study

Moreover, the studied countries can be assigned into the various clusters as illustrated in Table (5.12).

Table (5.12) Assigning countries to the resulting clusters for the hypothetical case study

Country	Cluster Number
Country (1)	1

Country (2)	1
Country (3)	3
Country (4)	4
Country (5)	5
Country (6)	6
Country (7)	4
Country (8)	5
Country (9)	1
Country (10)	5
Country (11)	4
Country (12)	4
Country (13)	3
Country (14)	4
Country (15)	1
Country (16)	5
Country (17)	2
Country (18)	2
Country (19)	4
Country (20)	4

According to the above stated outcomes, the countries that are similar in regard to the concerned location decision-making factors lie within the same cluster, while countries that are different from each other are included in different clusters. In fact, these findings would provide the entrepreneur who is keen to locate the entrepreneurial facility in some foreign markets that are characterized by the most fitting conditions for the new born business to fulfill the envisioned goals of its founder with an efficient tool to promote the selection process of the best international location to establish the entrepreneurial venture. Table (5.13) illustrates the similar countries in each of the resulting clusters.

Table (5.13) Groups of countries assigned to clusters for the hypothetical case study

Cluster	Countries
1	Country (1) Country (2) Country (9) Country (15)
2	Country (17) Country (18)
3	Country (3) Country (13)
4	Country (4) Country (7) Country (11) Country (12) Country (14) Country (19) Country (20)
5	Country (5) Country (8) Country (10) Country (16)
6	Country (6)

The improvement in the location decision-making process is primarily derived from restricting potential possible locations to accommodate the entrepreneurial facility into a limited number of clusters that consist of similar countries instead of the far larger pool of individual countries to compare, evaluate and then choose the best alternative among them. This amelioration also confirms that a valid good solution to the global facility location problem of

the small entrepreneurial enterprises could be obtained through applying the clustering data analysis algorithm.

Another advantage of implementing the developed clustering model is offering the decision maker a higher flexibility to select between available alternatives within the same cluster. Since each cluster includes countries that are similar in their attractiveness attributes, the entrepreneur can always have more options to establish the business in another country that belong to the same cluster in case of the inability to pursue the preferred choice due to reasons that did not exist when the list of potential countries was developed, such as political disturbances or natural disasters.

Moreover, the transformation of real values of the decision-making factors' numerical data into binary variables in the calculation of the JSC is also significant for defining the level of strength of these decision-making factors. This is important to identify the locations (countries) based on their similarities in including a strong level of particular decision-making factor(s). Therefore, countries could be joined together in distinct clusters depending on the similar strength level of the decision-making factor(s) they possess.

Therefore, in the previous case of the sub-factor, cost of business start-up procedures and after the conversion of its numerical data into binary variables, the explored countries can be grouped into four distinct clusters according to the strength level of that decision-making sub-factor.

Table (5.14) Countries assigned to clusters for the hypothetical case study based on the strength level of the decision-making sub-factor: cost of business start-up procedures

Cluster	Countries

(1) Cost of start-up procedures $0.2 - 2.45$	Country (1) Country (3) Country (4) Country (5) Country (6) Country (10) Country (11) Country (13) Country (15) Country (16)
(2) Cost of start-up procedures $2.45 - 4.7$	Country (2) Country (7) Country (17) Country (20)
(3) Cost of start-up procedures $4.7 - 6.95$	Country (8) Country (9) Country (19)
(4) Cost of start-up procedures $6.95 - 9.2$	Country (12) Country (14) Country (18) Country (20)

5.2 Real-world example

In the previous hypothetical case study the assumption was that all needed numerical data were available for all of the identified location decision-making factors. However, this is not always true where some of the numerical data for one or more factors of one or more countries are not available.

The validity of the introduced clustering analysis model can be also tested through applying the complete linkage-clustering algorithm (CLINK) on selected samples with real time data obtained from the World Bank's database. The first sample consists of the top 20 countries with the highest GDP for which most of the numerical values of the pre-defined decision-making factors are available and the similarity coefficient that can be used is the Euclidean distance.

5.2.1 Creating the list of investigated countries

The first step in applying the developed clustering analysis model is creating the list of elected countries to represent the objects for which the similarities and dissimilarities, in respect to attributes of the model that are represented by the specified location-attraction factors to entrepreneurs, are measured and then gathered in homogeneous groups or clusters.

Unlike the procedure explained for developing the list of countries in the previous hypothetical case study, the countries that will be included in the list for this real-world example are selected based on the completeness of numerical data within the World Bank's database of the decisive factors for better selecting a best-fit location to establish the entrepreneurial activity. In other words, any potential country that misses most of the numerical data of any decision-making factors in the World Bank's database will not be included in the list.

The countries that will be included on the list for this real-world example are the top twenty countries with the highest GDP (the G20). The GDP indicator is considered because it is a measure of the size of a nation's economy and it measures the buying power of a nation over a given time period. Moreover, GDP is also used as an indicator of a nation's overall standard of living because, generally, a nation's standard of living increases as GDP increases.

Depending on the indicated conditions, the considered list of countries in the real-world example is given in Table (5.15).

Table (5.15) Final list of the G20 countries for the real-world example

#	Country
1	United States
2	China
3	Japan
4	Germany
5	United Kingdom
6	France
7	Brazil
8	Italy
9	India
10	Russian Federation
11	Canada
12	Australia
13	Korea, Rep.
14	Spain
15	Mexico
16	Indonesia
17	Netherlands
18	Turkey

19	Saudi Arabia
20	Sweden

5.2.2 Specifying the decision making factors

The second step in executing the model with the selected sample is similar to the procedure of the hypothetical case study. However, the set of the considered decision-making factors that has been developed for the hypothetical case study cannot be applied directly to the real-world sample case due to the considerable unavailable data of the factors related to infrastructure. Therefore, the complete set of the considered decision-making factors will be modified and updated by taking out the related infrastructure factors and it is indicated in following table.

Table (5.16) The updated list of location decision-making factors and associated sub-factors for the real-world example

Main factor	Decision-making sub-factors (attributes)
Business start-up cost and procedure	Cost of business start-up procedures Start-up procedures to register a business Time required to start a business Patent applications Trademark applications Charges for the use of intellectual property
Financing small and medium enterprises	Firms using banks to finance investment Lending interest rate Foreign direct investment
Tax rates and structure	Total tax rate Profit tax Taxes on goods and services

Governmental regulations and policies	Exports of goods and services
	Trade in services
	Net official development assistance and official aid received
Labor and skills	Labor force with tertiary education
	Secondary education, vocational pupils
	Government expenditure on education
	Wage and salaried workers
	Unemployment
Technology advancement	High-technology exports
	Internet users
	Fixed broadband subscriptions
	Research and development expenditure
	Researchers in research and development
	Manufacturing, value added
Competition	Listed domestic companies

5.2.3 Collecting and setting up data

The needed data that represent the decision-making factors are gathered from the World Bank's database. The numerical values for each of the sub-factors for the main decision-making factors are shown in the Appendix.

5.2.4 Assigning weights to data of decision-making factors

As mentioned in section 4.2, weights might be assigned to each identified location decision factor based on the degree of importance it has been given in the literature or on how strongly entrepreneurship scholars emphasized it in their research. Assignment of weights to the decision-making factors helps to find out mismatches on expectations. The assignment of weights also helps decision makers to be less defensive and be more objective in evaluating the available alternatives.

Since the list of the decision-making factors for the real-world example has been updated as discussed in section 5.2.2, the assigned weights must be also updated. The updated assigned weights for each of the decision-making sub-factors are shown in Table (5.17).

Table (5.17) The updated weights assigned to the location decision factors for the real-world example

#	Decision-making factors (attributes)	Weight (%)
1	Cost of business start-up procedures	1.56
2	Start-up procedures to register a business	0.9
3	Time required to start a business	0.8
4	Patent applications	0.7
5	Trademark applications	0.6
6	Charges for the use of intellectual property	0.5
7	Firms using banks to finance investment	5
8	Lending interest rate	8
9	Foreign direct investment	6
10	Total tax rate	4
11	Profit tax	3
12	Taxes on goods and services	2
13	Exports of goods and services	0.4
14	Trade in services	0.3
15	Net official development assistance and official aid received	0.4
16	Labor force with tertiary education	0.09
17	Secondary education, vocational pupils	0.08
18	Government expenditure on education	0.1
19	Wage and salaried workers	0.07
20	Unemployment	0.06
21	High-technology exports	10.4
22	Internet users	9
23	Fixed broadband subscriptions	9

24	Research and development expenditure	14
25	Researchers in R&D	11
26	Manufacturing, value added	12
27	Listed domestic companies	0.04

5.2.5 Implementing the clustering analysis model

After collecting and setting up the required data to be installed in the developed model, the complete linkage clustering method (CLINK) with Euclidean distance coefficient is applied.

Implementation of the developed clustering model will form clusters consist of homogeneous groups combining countries that are most similar in respect to the location decision-making factors.

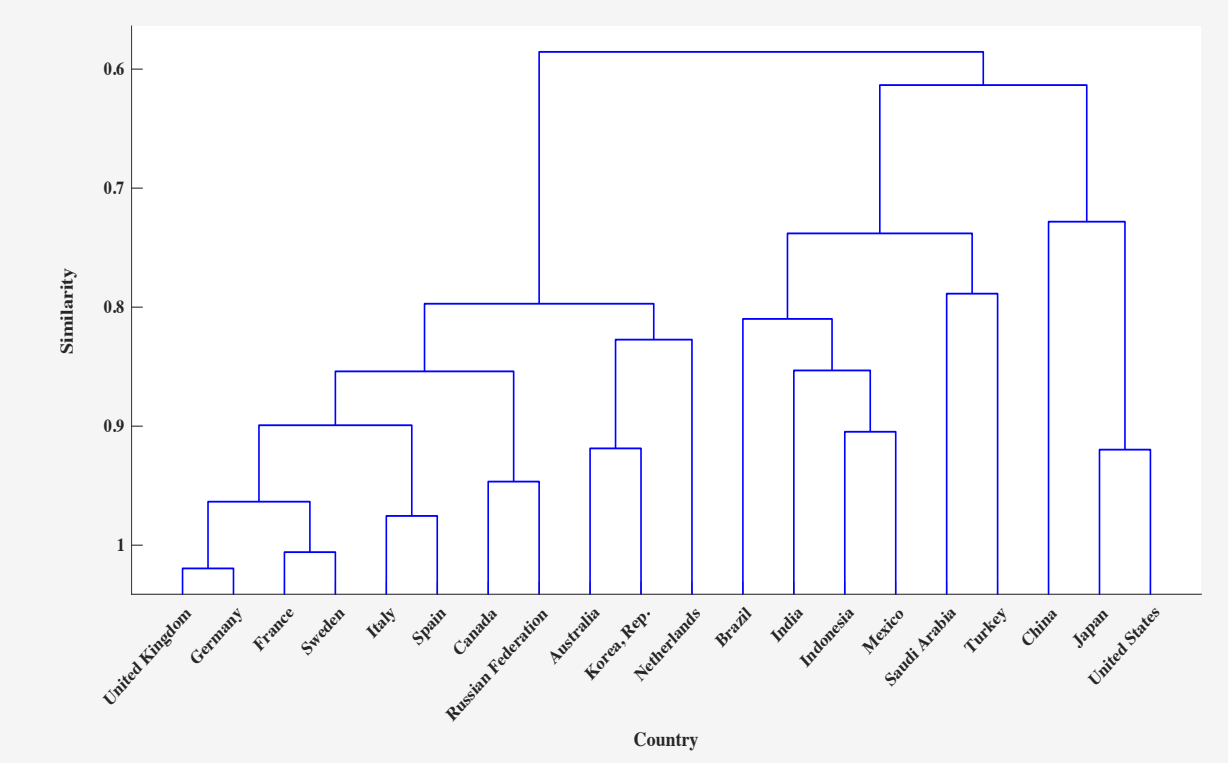

Figure (5.2) Dendrogram of the developed model for the real-world example using Euclidean distance with complete linkage clustering

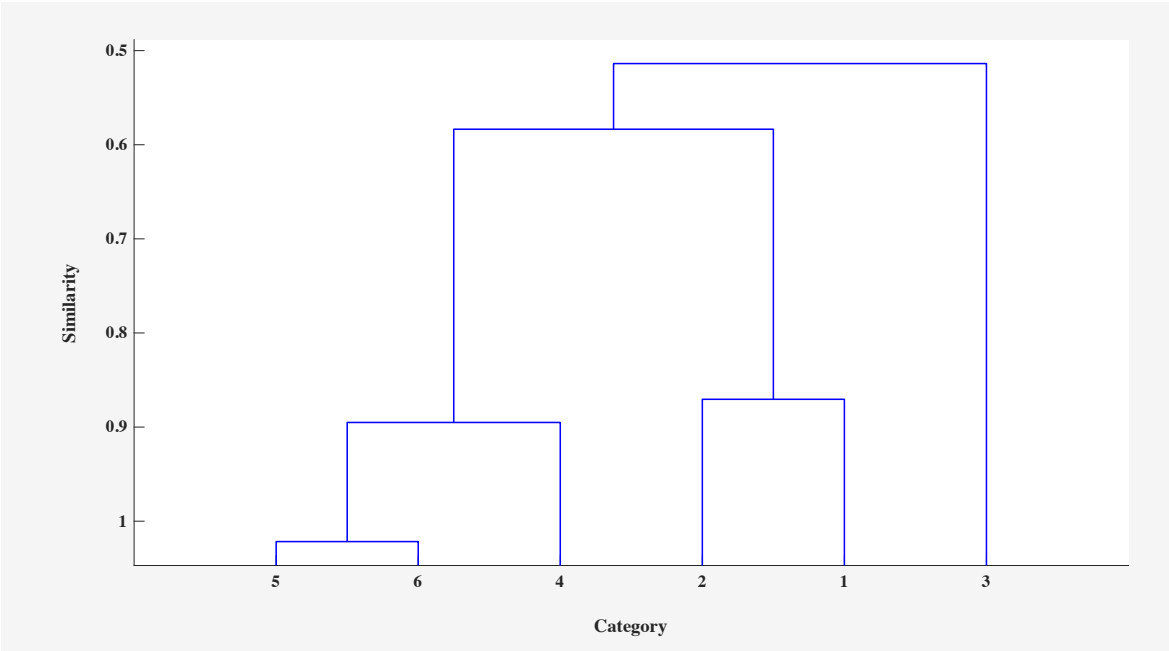

Figure (5.3) Dendrogram of clustering the real-world example countries using Euclidean distance with complete linkage clustering in six categories

Therefore, the investigated countries can be assigned into six distinctive clusters as indicated in Table (5.18).

Table (5.18) Assigning the G20 countries to clusters for the real-world example

Country	Cluster Number
United States	1
China	2
Japan	1
Germany	3
United Kingdom	3

France	3
Brazil	4
Italy	3
India	4
Russian Federation	3
Canada	3
Australia	3
Korea, Rep.	3
Spain	3
Mexico	4
Indonesia	4
Netherlands	3
Turkey	6
Saudi Arabia	5
Sweden	3

Moreover, Table (5.19) below conveys how the considered countries are distributed among the resulting clusters.

Table (5.19) Distribution of countries among clusters for the real-world example

Cluster	**Countries**

1	United States Japan
2	China
3	Germany United Kingdom France Italy Russian Federation Canada Australia Korea, Rep. Spain Netherlands Sweden
4	Brazil India Mexico Indonesia
5	Saudi Arabia
6	Turkey

The resulting clustering trend occurs because there are other decision-making factors that are most likely affecting the attractiveness of locations to entrepreneurs who seek to start up their ventures internationally. Furthermore, the results of the developed model emphasize the impact

of pre-defined location decision-making factors on the process of selecting the best-fit location for the entrepreneurial firms. These results also prove to a further extent the validity of the developed clustering analysis model, as well as how heavily the global location decision-making process for small entrepreneurial businesses is affected by attractive factors to entrepreneurs that characterize the studied potential locations.

Moreover, ranking countries within each cluster might add more value to some interested entrepreneurs. In this research the ranking is conducted by comparing the total values of the weighted decision making factors for the investigated countries. The larger the total value of a country, the higher the rank of that country. The total value of the weighted decision making factors for a country *i* is given by:

$$\sum_{a_i=1}^{n} a_i \times (weight\ of\ a_i)$$

where a_i is the normalized numerical value of a decision-making factor for country i and n is the total number existing in the i country.

The resulting rank shall be considered as initial ranking: making more reliable decisions requires a deeper investigation of the attractiveness factors for entrepreneurial firms that exist in each of these countries.

Table (5.20) Ranks of countries among each cluster for the real-world example

Cluster	Countries	Rank
1	Japan	1
	United States	2
2	China	1
3	Korea, Rep.	1

	Germany	2
	Sweden	3
	Netherlands	4
	France	5
	United Kingdom	6
	Australia	7
	Canada	8
	Spain	9
	Italy	10
	Russian Federation	11
4	Brazil	1
	Mexico	2
	India	3
	Indonesia	4
5	Saudi Arabia	1
6	Turkey	1

An overall ranking can be also obtained based on the decision-making factors that are considered in the research to put the G20 countries in a descending order to their attractiveness to entrepreneurship activities.

Table (5.21) Overall rank of the G20 countries for entrepreneurship in the real-world example

Country	**Rank**	**Cluster**
Korea, Rep.	1	3
Japan	2	1
United States	3	1
Germany	4	3
Sweden	5	3
China	6	2
Netherlands	7	3

France	8	3
United Kingdom.	9	3
Australia	10	3
Canada	11	3
Spain	12	3
Italy	13	3
Russian Federation	14	3
Brazil	15	4
Turkey	16	6
Mexico	17	4
India	18	4
Indonesia	19	4
Saudi Arabia	20	5

5.3 The effect of the number of identified location decision-making factors

The efficiency of the developed model is proportional to the number of location decision factors included in the process. Increasing the number of these factors would most likely result in generating more defined clusters. To examine the affected efficiency of the model by the increment of the number of decisive factors, two steps are carried out for the top 20 countries with the highest GDPs (the G20 countries).

Step one is developing clusters for the top 20 countries with the highest GDPs using only three decision-making factors. Data for three decision-making factors that are derived from the World Bank's database are shown in the table below.

Table (5.22) Data of three decision-making factors for the G20 countries

Country	Time required to start a business	Patent applications	Taxes on goods and services
United States	5.60	293706	0.60

China	31.40	127042	7.75
Japan	10.20	60030	5.05
Germany	12.50	17811	7.65
United Kingdom	5.25	7844	13.3
France	4.25	2033	10.95
Brazil	83.3	25683	7.65
Italy	5.75	781	10.30
India	31.50	30814	3.80
Russian Federation	10.85	16236	7.10
Canada	3.50	31283	2.70
Australia	2.50	23968	6.35
Korea, Rep.	4	46219	6.25
Spain	14	225	8.05
Mexico	6.30	14889	0
Indonesia	50.15	7321	5.65
Netherlands	4	288	11.35
Turkey	7.50	331	17.45
Saudi Arabia	19.75	135	0
Sweden	11.50	441	14.25

By processing the CLINK algorithm embedded in the developed model, the investigated countries would be assigned as homogenous groups into various distinctive clusters. The following dendrogram illustrates the groups of the top 20 countries with the highest GDPs using three decision-making factors.

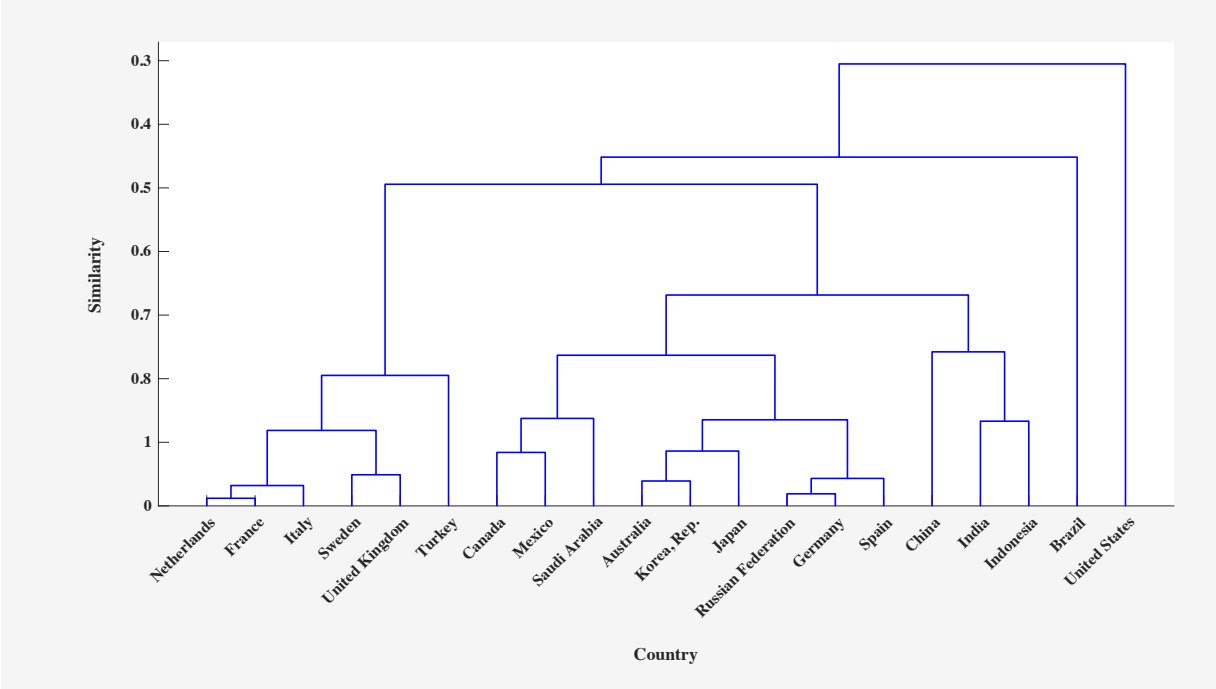

Figure (5.4) Dendrogram of clustering the G20 countries based on three decision-making factors

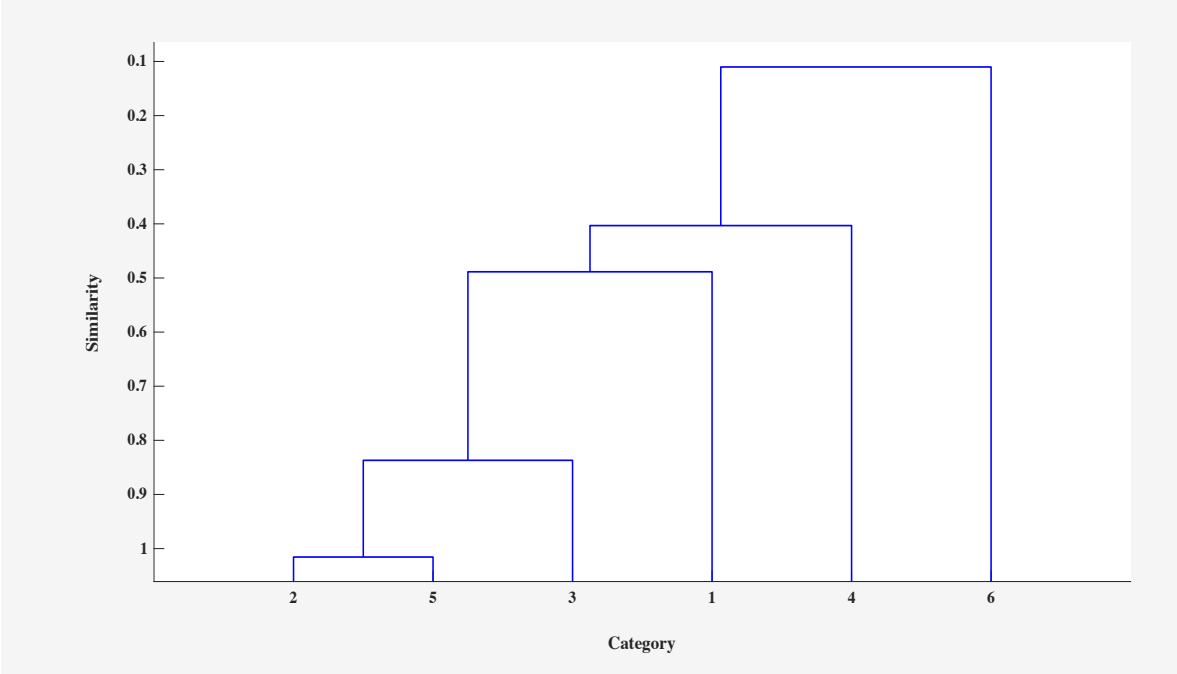

Figure (5.5) Dendrogram of clustering the G20 countries based on three decision-making factors in six categories

Therefore, the G20 countries are assigned into six clusters as shown in Table (5.23).

Table (5.23) Assigning the G20 countries to clusters based on three decision-making factors

Cluster #	Countries
1	United Kingdom France Italy Netherlands Turkey Sweden
2	China
3	Japan Germany Russian Federation Canada Australia Korea, Rep. Spain Mexico Saudi Arabia
4	Brazil
5	India Indonesia
6	United States

Step two is developing clusters for the top 20 countries with the highest GDPs after adding the data of three more decision-making factors. The data of the added three decision-making factors are also derived from the World Bank's database. Data of the added three decision-making factors for the G20 countries are given in Table (5.24).

Table (5.24) Data of the added three decision-making factors for the G20 countries

Country	**Start-up procedures to register a business**	**High-technology exports**	**Total tax rate**
United States	6	154353963992	43.90
China	11	559332162922.5	67.80
Japan	8	91529336519	51.30
Germany	9	184283164631	48.80
United Kingdom	6	69340644491	32
France	5	132183573785	62.70
Brazil	11.30	8848309553	69.20
Italy	5	26955337473	64.80
India	13.40	13750546786	60.60
Russian Federation	4.40	9249223001.5	47
Canada	2	26268767511	21.10
Australia	3	4237456601	47.60
Korea, Rep.	3	131953914182	33.20

Spain	7	14240904065	50
Mexico	6	45780911356	51.70
Indonesia	13	4899457279	29.70
Netherlands	4	69673950438.5	41
Turkey	8	2323079468	40.90
Saudi Arabia	12	272788564	15
Sweden	3	14933994823	49.10

By following the same procedure for the updated set of decision-making factors, different results are obtained. A dendrogram of outcomes for the updated set of six decision-making factors for the G20 countries is shown in Figure (5.6).

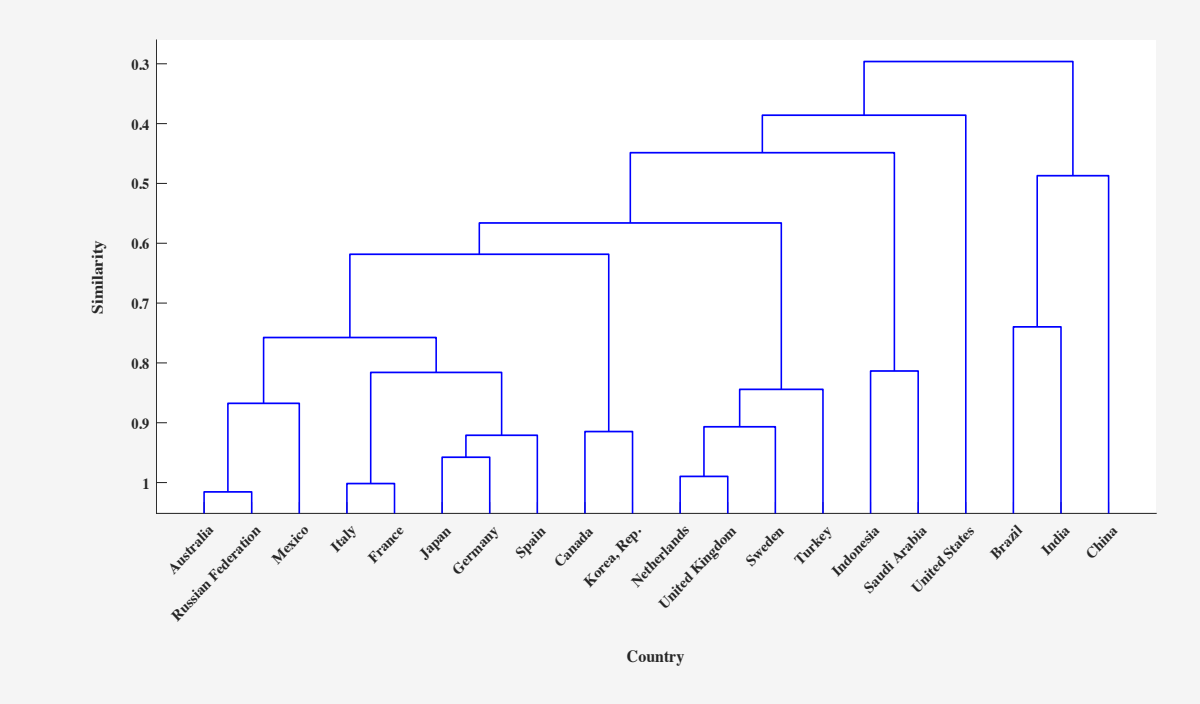

Figure (5.6) Dendrogram of clustering the G20 countries based on six decision-making factors

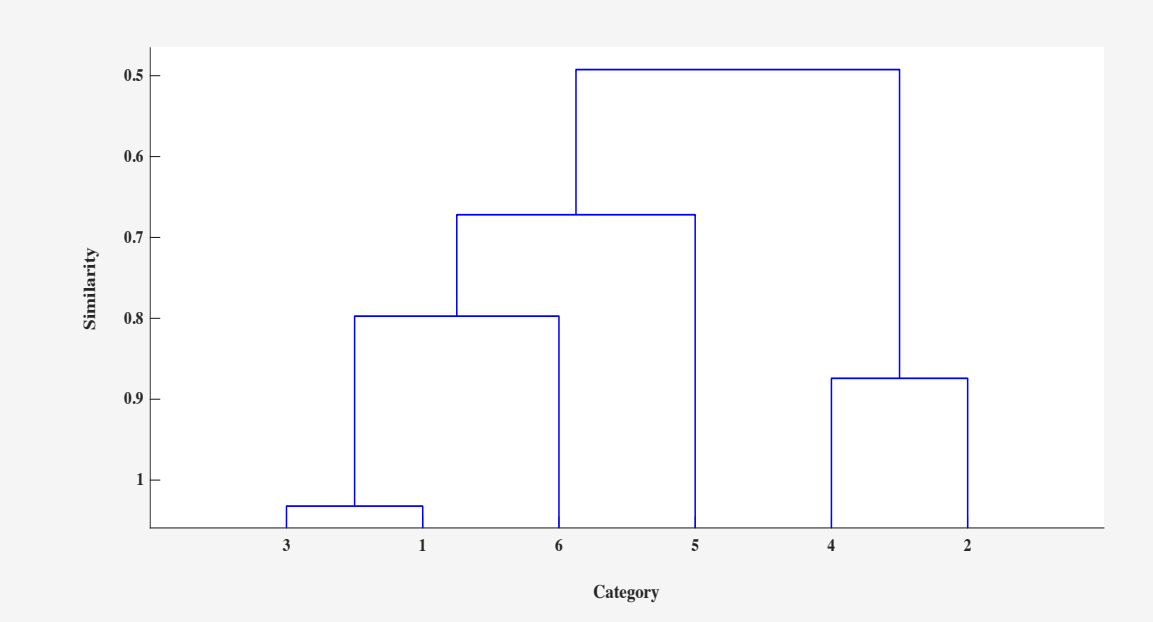

Figure (5.7) Dendrogram of clustering the G20 countries based on six decision-making factors in six categories

The addition of more decision-making factors to the developed model results in different assignments of the considered countries into the newly formed clusters. The yielded clusters and assigned countries to each cluster are given in the Table (5.25).

Table (5.25) Assigning the G20 countries to clusters based on six decision-making factors

Cluster #	**Countries**
1	United Kingdom Netherlands Turkey Sweden
2	China

3	Japan Germany France Italy Russian Federation Canada Australia Korea, Rep. Spain Mexico
4	Brazil India
5	United States
6	Indonesia Saudi Arabia

Comparing the resulting clusters from the two previous steps indicates that the inclusion of more location decision factors in the process of the developed model leads to generating different sets of clusters and some of the studied countries in the step one are assigned to different clusters in the step two. This clearly shows that the number of decision-making factors under consideration affects the proposed model.

Moreover, it is most likely expected that the outcomes of the model keep progressing as more location decision factors are added.

5.4 The effect of assigning weights to the location decision-making factors

In this research, weights are assigned to the identified location decision-making factors based on the degree of importance each of them has been given by the scholars and researchers of entrepreneurship in selecting the best-fit location for the entrepreneurial ventures. The existence of these factors is considered essential to the success of entrepreneurship in any potential location. However, in many cases different entrepreneurs are interested in some or most of the location decision-making factors with different degrees of importance due to the nature and type of their business, which requires adjustment of their given weights accordingly.

To test the effects of assigned weights on the introduced clustering model, the weights assigned to the location decision-making factors in Table (5.16) are going to be modified according to the need of the assumed specific type of business: then the model will be applied in the real-world example of the G20 countries.

Assuming that the considered business requires a highly educated work force, the updated list of the decisive factors and their weights are given in the table below.

Table (5.26) The updated weights assigned to location decision-making factors for an assumed technological small venture

#	Decision-making factors (attributes)	Weight (%)
1	Cost of business start-up procedures	1.56
2	Start-up procedures to register a business	0.9
3	Time required to start a business	0.9
4	Patent applications	0.7
5	Trademark applications	0.6

6	Charges for the use of intellectual property	0.5
7	Firms using banks to finance investment	0.07
8	Lending interest rate	0.09
9	Foreign direct investment	0.08
10	Total tax rate	4
11	Profit tax	0.06
12	Taxes on goods and services	0.2
13	Exports of goods and services	0.4
14	Trade in services	0.3
15	Net official development assistance and official aid received	0.2
16	Labor force with tertiary education	14
17	Secondary education, vocational pupils	6
18	Government expenditure on education	2
19	Wage and salaried workers	11
20	Unemployment	3
21	High-technology exports	4.4
22	Internet users	9
23	Fixed broadband subscriptions	9
24	Research and development expenditure	14
25	Researchers in R&D	11
26	Manufacturing, value added	6
27	Listed domestic companies	0.04

By applying the developed model into the data provided for the G20 in the Appendix, the resulting clusters of countries can be obtained in the following dendrograms.

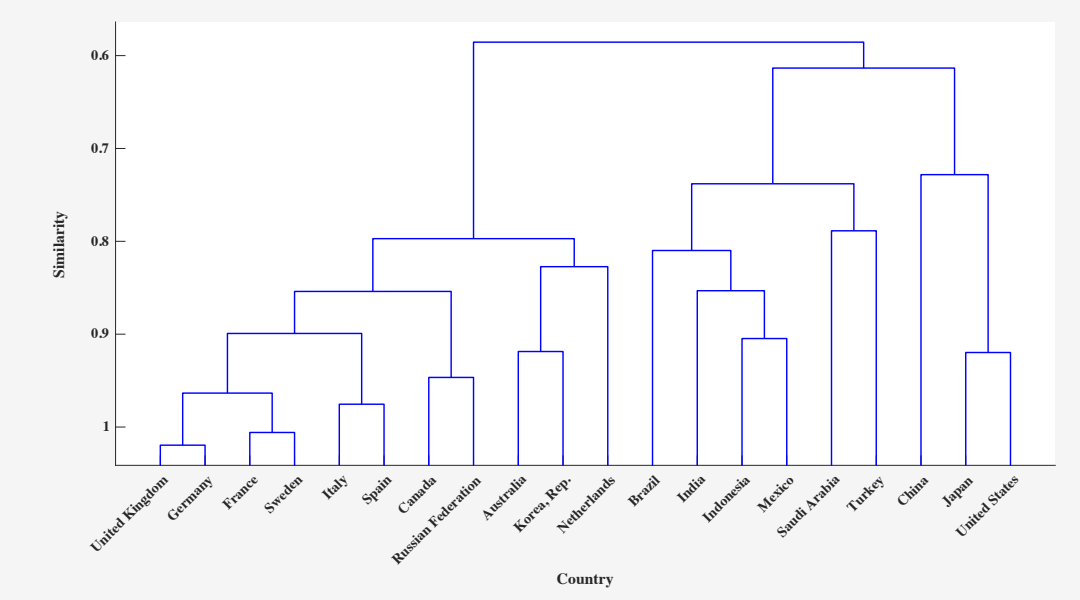

Figure (5.8) Dendrogram of the developed model for the modified real-world example using Euclidean distance with complete linkage clustering for a business that requires highly educated work force

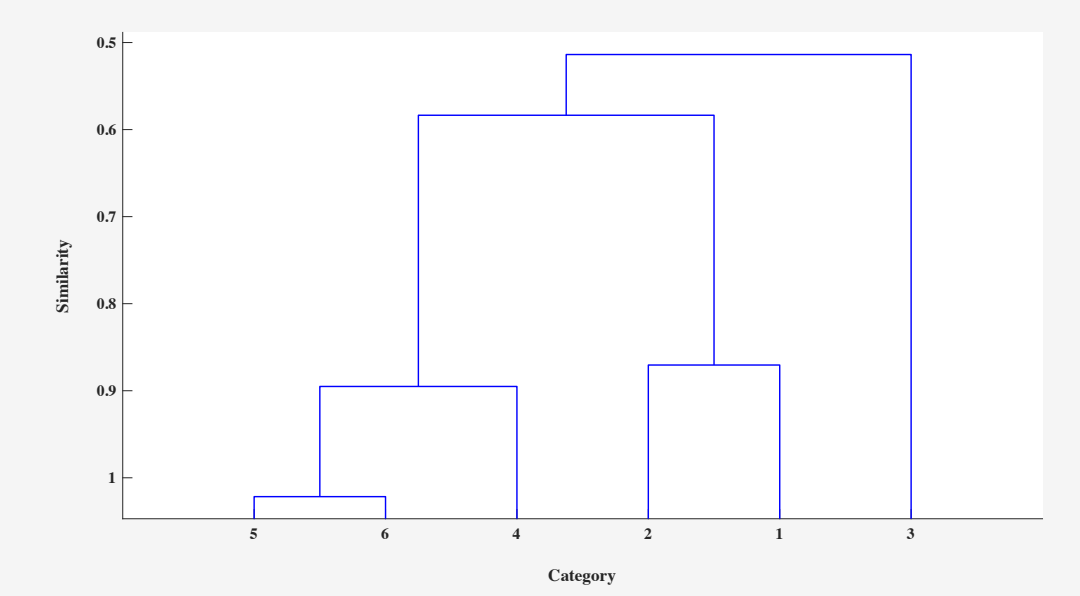

Figure (5.9) Dendrogram of clustering the modified real-world example countries using Euclidean distance with complete linkage clustering in six categories

Table (5.27) Distribution of countries among clusters for the modified real-world example

Cluster	Countries
1	United States Japan
2	China
3	Germany United Kingdom France Italy Russian Federation Canada Australia Korea, Rep. Spain Netherlands Sweden
4	Brazil India Mexico Indonesia
5	Saudi Arabia
6	Turkey

The change in assigned weights to decision-making factors does not affect the countries that are included in each cluster. However, the rank of countries is highly influenced by the change in assigned weights to decision-making factors. In fact, this also leads to the change in the ranking of the countries among each individual cluster.

Table (5.28) Ranks of countries among each cluster for the modified real-world example

Cluster	Countries	Rank
1	United States	1
	Japan	2
2	China	1
3	Sweden	1
	Germany	2
	France	3
	Netherlands	4
	Canada	5
	Korea, Rep.	6
	United Kingdom	7
	Spain	8
	Russian Federation	9
	Australia	10
	Italy	11
4	Brazil	1
	Mexico	2
	India	3
	Indonesia	4
5	Saudi Arabia	1
6	Turkey	1

Moreover, the overall rank of countries is shown in the following table.

Table (5.29) Overall rank of the G20 countries for entrepreneurship in the modified real-world example

Country	Rank	Cluster
Sweden	1	3
United States	2	1
Germany	3	3
France	4	3
Netherlands	5	3
Canada	6	3
Korea, Rep.	7	3
United Kingdom	8	3
Japan	9	1
China	10	2
Spain	11	3
Russian Federation	12	3
Australia	13	3
Italy	14	3
Brazil	15	4
Mexico	16	4
Turkey	17	6
Saudi Arabia	18	5
India	19	4
Indonesia	20	4

5.5 Applying the model into a large size real-world sample

One prominent advantage of the developed model is its flexibility. The flexibility of the proposed similarity coefficient-based approaches is categorized into two levels: (1) the model is

flexible in its application into either limited or large and complex decision-making problems, and (2) it is also highly flexible when adding, removing or editing the decision-making factor being considered.

Both the validity and the flexibility of the model can be tested through applying the complete linkage-clustering algorithm with the Euclidean distance coefficient into a large size sample with real time data. The sample consists of the top 100 countries based on their average rank applied on the weighted numerical data and comparing them with the most credible indices.

Furthermore, the flexibility of the proposed clustering analysis model will be also examined by applying several clustering analysis approaches: i.e., several similarity coefficients with various clustering algorithms will be applied on the same large size real-world sample.

5.5.1 Application of the developed model into the large size real-world sample

In this section the complete linkage-clustering algorithm (CLINK) with the Euclidean distance similarity coefficient (as the proposed model in the research) is going to be applied into the large size real-world sample.

5.5.1.1 Creating the list of investigated countries

The list of countries that will be investigated in the large size real-world sample consists of one hundred countries. The countries will be selected based on their entrepreneurial attractiveness level which is derived from average rank applied on the weighted numerical data. Table (5.30) below illustrates the top 100 investigated countries that are included in the final list (the G20 countries as well as the rest of countries alphabetically).

Table (5.30) Final list of the countries for the large size real-world sample

#	Country	#	Country	#	Country
1	United States	35	Chile	69	Moldova
2	China	36	Colombia	70	Montenegro
3	Japan	37	Costa Rica	71	Morocco
4	Germany	38	Croatia	72	Namibia
5	United Kingdom	39	Cyprus	73	Nigeria
6	France	40	Czech Republic	74	Norway
7	Brazil	41	Denmark	75	Oman
8	Italy	42	Dominican Republic	76	Panama
9	India	43	Ecuador	77	Peru
10	Russian Federation	44	Egypt, Arab Rep.	78	Philippines
11	Canada	45	El Salvador	79	Poland
12	Australia	46	Estonia	80	Portugal
13	Korea, Rep.	47	Finland	81	Puerto Rico
14	Spain	48	Gabon	82	Qatar
15	Mexico	49	Georgia	83	Romania
16	Indonesia	50	Ghana	84	Serbia
17	Netherlands	51	Greece	85	Singapore
18	Turkey	52	Hong Kong SAR, China	86	Slovak Republic
19	Saudi Arabia	53	Hungary	87	Slovenia
20	Sweden	54	Iceland	88	South Africa
21	Albania	55	Iran, Islamic Rep.	89	Sri Lanka
22	Algeria	56	Ireland	90	Swaziland

23	Argentina	57	Israel	91	Switzerland
24	Armenia	58	Jamaica	92	Tajikistan
25	Austria	59	Jordan	93	Thailand
26	Azerbaijan	60	Kazakhstan	94	Trinidad and Tobago
27	Bahrain	61	Kuwait	95	Tunisia
28	Barbados	62	Kyrgyz Republic	96	Ukraine
29	Belgium	63	Latvia	97	United Arab Emirates
30	Bolivia	64	Lebanon	98	Uruguay
31	Bosnia and Herzegovina	65	Lithuania	99	Vietnam
32	Botswana	66	Luxembourg	100	Zambia
33	Brunei Darussalam	67	Macedonia, FYR		
34	Bulgaria	68	Malaysia		

5.5.1.2 Specifying the decision making factors

As applied in the previous two examples, the set of decision-making factors that has been previously developed and listed in Table (5.15) is going to be used to group the listed countries based on their similarities and dissimilarities.

5.5.1.3 Collecting and setting up data

As in the last real-world example, the data from the World Bank's database will be used to represent the decision-making factors and relate them to the countries. The numerical values for each of the sub-factors are given in the Appendix.

5.5.1.4 Assigning weights to the data of the decision-making factors

The weights that have been listed in Table (5.16) will be assigned to the decision-making factors.

5.5.1.5 Implementing the clustering analysis model

The next step is to apply the developed model to the large size real-world sample; here the selected clustering method is the complete linkage (CLINK) with Euclidean distance for the similarity coefficient.

Similar to the results obtained in the previous example, distinct clusters of the considered countries will be obtained. The formed clusters are shown in the dendrograms shown below

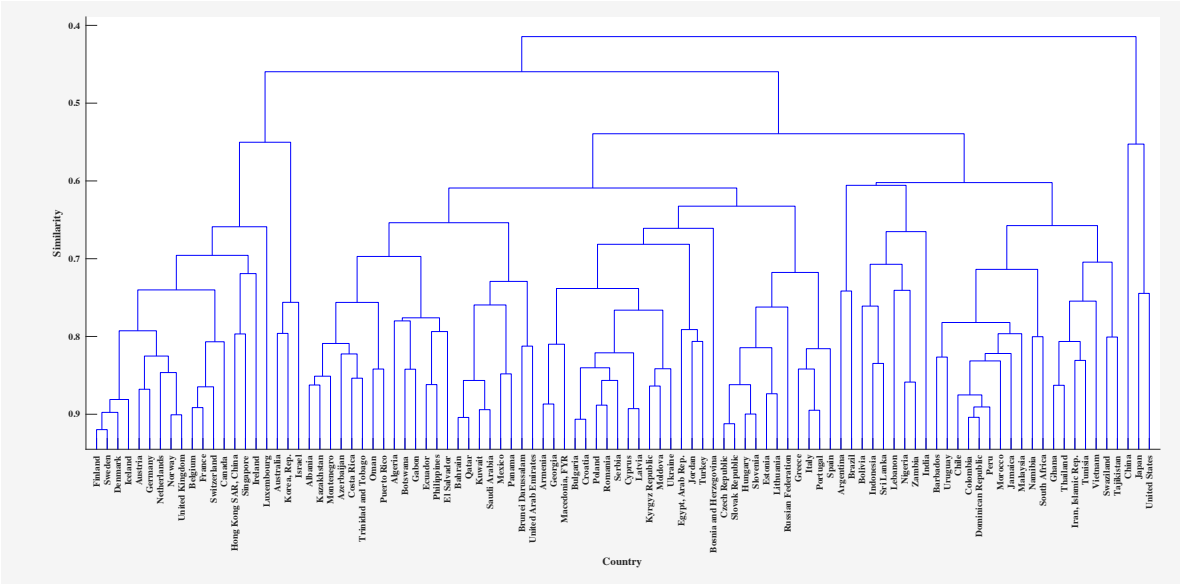

Figure (5.10) Dendrogram of clustering the large size real-world sample countries using Euclidean distance with complete linkage clustering

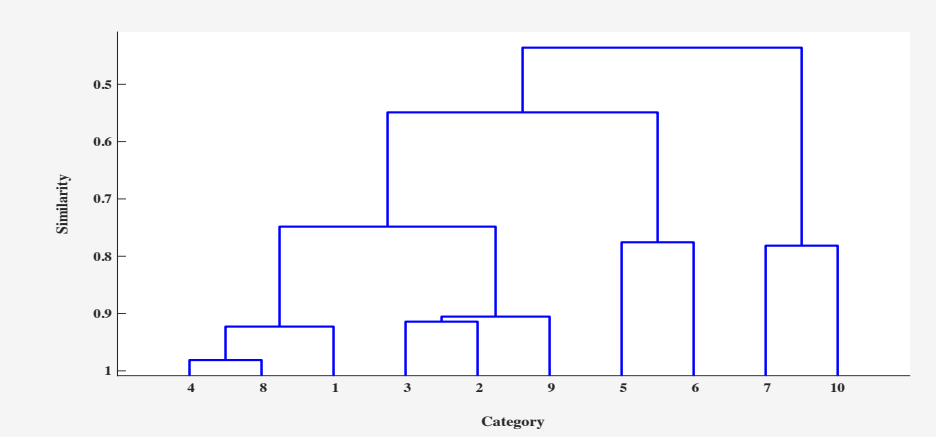

Figure (5.11) Dendrogram of clustering the large size real-world sample countries using Euclidean distance with complete linkage clustering in ten categories

Table (5.31) Assigning the countries to clusters using Euclidean distance with complete linkage clustering in ten categories

Cluster #	Countries	Cluster #	Countries
1	United States Japan	2	China
3	Korea, Rep. Australia Israel	4	Brazil Argentina
5	Germany Austria Belgium Canada Denmark Finland France Hong Kong SAR, China Iceland Ireland Luxembourg Netherlands Norway Singapore Sweden Switzerland United Kingdom	6	Italy Russian Federation Spain Portugal Czech Republic Estonia Greece Hungary Lithuania Slovak Republic Slovenia
7	India	8	Turkey

	Bolivia		Armenia
	Indonesia		Bosnia and Herzegovina
	Lebanon		Kyrgyz Republic
	Nigeria		Poland
	Sri Lanka		Bulgaria
	Zambia		Ukraine
			Croatia
			Macedonia, FYR
			Serbia
			Latvia
			Moldova
			Georgia
			Romania
			Cyprus
			Jordan
			Egypt, Arab Rep.
9	Saudi Arabia	10	Malaysia
	United Arab Emirates		Thailand
	Qatar		Iran, Islamic Rep.
	El Salvador		Tajikistan
	Kuwait		Vietnam
	Oman		Barbados
	Bahrain		Chile
	Brunei Darussalam		Colombia
	Mexico		Dominican Republic
	Montenegro		Jamaica
	Philippines		Peru
	Albania		Uruguay

Azerbaijan	Swaziland
Kazakhstan	South Africa
Algeria	Morocco
Botswana	Tunisia
Gabon	Ghana
Trinidad and Tobago	Namibia
Costa Rica	
Ecuador	
Panama	
Puerto Rico	

Furthermore, the obtained ranking of countries within each cluster is shown in the following table.

Table (5.32) Ranking of countries among each cluster for the large size real-world sample

Cluster	Countries	Rank
1	United States	1
	Japan	2
2	China	1
3	Korea, Rep.	1
	Australia	2
	Israel	3
4	Argentina	1
	Brazil	2
5	Germany	1
	Denmark	2
	Sweden	3
	Finland	4
	Switzerland	5

	Netherlands	6
	Iceland	7
	Norway	8
	Austria	9
	France	10
	United Kingdom	11
	Belgium	12
	Singapore	13
	Luxembourg	14
	Canada	15
	Ireland	16
	Hong Kong SAR, China	17
6	Czech Republic	1
	Slovenia	2
	Estonia	3
	Hungary	4
	Slovak Republic	5
	Spain	6
	Italy	7
	Russian Federation	8
	Portugal	9
	Lithuania	10
	Greece	11
7	Lebanon	1
	India	2
	Bolivia	3
	Indonesia	4
	Sri Lanka	5
	Nigeria	6

	Zambia	7
	Poland	1
	Turkey	2
	Ukraine	3
	Serbia	4
	Latvia	5
	Croatia	6
	Bulgaria	7
	Bosnia and Herzegovina	8
8	Romania	9
	Moldova	10
	Macedonia, FYR	11
	Cyprus	12
	Armenia	13
	Georgia	14
	Jordan	15
	Egypt, Arab Rep.	16
	Kyrgyz Republic	17
	Azerbaijan	1
	Costa Rica	2
	Trinidad and Tobago	3
	Puerto Rico	4
	Montenegro	5
9	Mexico	6
	Bahrain	7
	Qatar	8
	Kazakhstan	9
	United Arab Emirates	10
	Ecuador	11

	Albania	12
	Philippines	13
	Botswana	14
	Oman	15
	El Salvador	16
	Kuwait	17
	Brunei Darussalam	18
	Panama	19
	Saudi Arabia	20
	Algeria	21
	Gabon	22
10	Malaysia	1
	Barbados	2
	Uruguay	3
	Chile	4
	Colombia	5
	Morocco	6
	Dominican Republic	7
	Tunisia	8
	Peru	9
	Thailand	10
	South Africa	11
	Jamaica	12
	Vietnam	13
	Tajikistan	14
	Iran, Islamic Rep.	15
	Namibia	16
	Ghana	17
	Swaziland	18

The overall obtained ranking based on the considered decision-making factors is listed in the table below.

Table (5.33) Overall rank of the top countries for entrepreneurship in the large size real-world sample

Country	Rank	Country	Rank
Korea, Rep.	1	Bosnia and Herzegovina	51
United States	2	Trinidad and Tobago	52
Japan	3	Romania	53
Israel	4	Colombia	54
Germany	5	Morocco	55
Denmark	6	Moldova	56
Sweden	7	Puerto Rico	57
China	8	Macedonia, FYR	58
Finland	9	Montenegro	59
Switzerland	10	Cyprus	60
Netherlands	11	Dominican Republic	61
Iceland	12	Tunisia	62
Norway	13	Peru	63
Austria	14	Armenia	64
France	15	Thailand	65
United Kingdom	16	Mexico	66
Belgium	17	Georgia	67
Singapore	18	South Africa	68
Czech Republic	19	Jamaica	69
Slovenia	20	Bahrain	70
Australia	21	Qatar	71
Luxembourg	22	Vietnam	72
Canada	23	Kazakhstan	73

Estonia	24	Jordan	74
Ireland	25	India	75
Hungary	26	United Arab Emirates	76
Slovak Republic	27	Egypt, Arab Rep.	77
Spain	28	Bolivia	78
Argentina	29	Ecuador	79
Italy	30	Albania	80
Russian Federation	31	Indonesia	81
Portugal	32	Philippines	82
Hong Kong SAR, China	33	Tajikistan	83
Malaysia	34	Botswana	84
Lithuania	35	Kyrgyz Republic	85
Greece	36	Oman	86
Brazil	37	Iran, Islamic Rep.	87
Poland	38	Sri Lanka	88
Barbados	39	El Salvador	89
Uruguay	40	Kuwait	90
Turkey	41	Namibia	91
Ukraine	42	Brunei Darussalam	92
Serbia	43	Panama	93
Latvia	44	Saudi Arabia	94
Croatia	45	Algeria	95
Lebanon	46	Nigeria	96
Bulgaria	47	Ghana	97
Azerbaijan	48	Swaziland	98
Chile	49	Zambia	99
Costa Rica	50	Gabon	100

5.5.2 Application of other clustering analysis approaches into the large size real-world sample

The validity and flexibility of the developed model can be also tested through applying different clustering algorithms in order to understand the different or similar effects these clustering algorithms have on the considered data in forming the desired clusters.

To do so four different approaches are applied into the large size real-world data:

Approach 1: Applying Euclidean distance with complete linkage clustering.

Approach 2: Applying Euclidean distance with average linkage clustering.

Approach 3: Applying CityBlock with complete linkage clustering.

Approach 4: Applying CityBlock with average linkage clustering.

5.5.2.1 Approach 1: Applying Euclidean distance with complete linkage clustering

This approach has been discussed in the last section as the developed clustering model.

5.5.2.2 Approach 2: Applying Euclidean distance with average linkage clustering

The application of Euclidean distance will result in forming clusters that are illustrated in the following dendrograms.

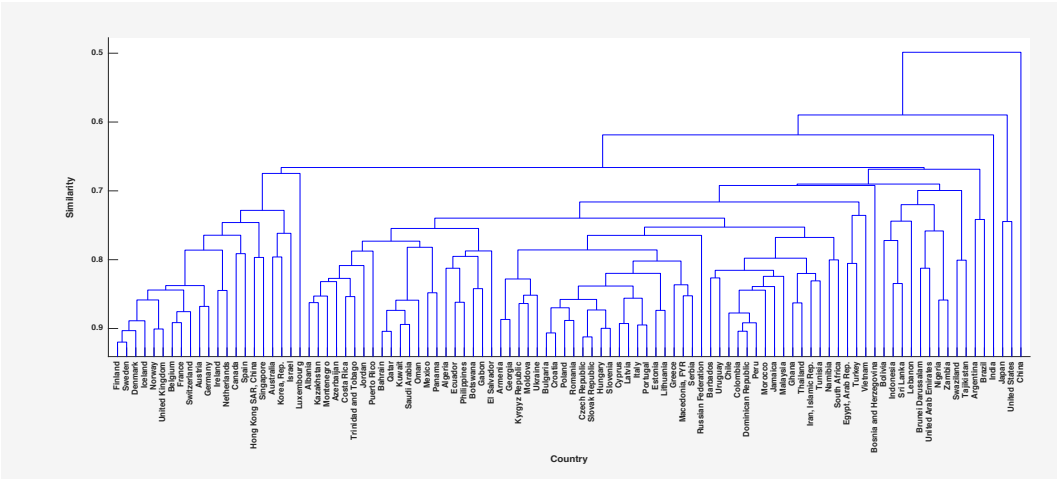

Figure (5.12) Dendrogram of clustering the large size real-world sample countries using Euclidean distance with average linkage clustering

Also, the dendrogram of clustering the countries using Euclidean distance with average linkage clustering in ten categories is given below.

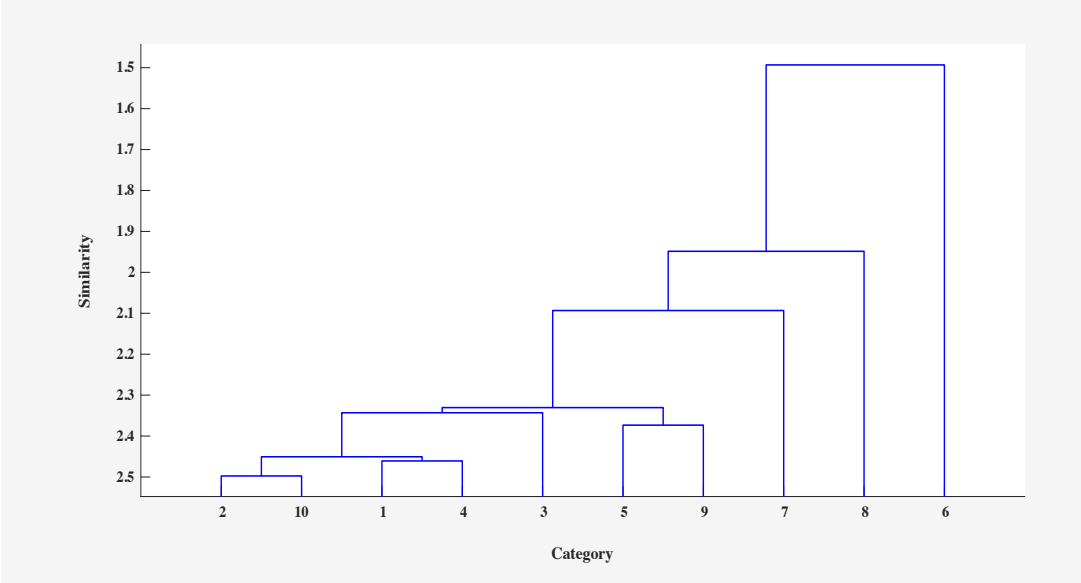

Figure (5.13) Dendrogram of clustering the large size real-world sample countries using Euclidean distance with average linkage clustering in ten categories

Table (5.34) Assigning the countries to clusters using Euclidean distance with average linkage clustering in ten categories

Cluster #	**Countries**		**Cluster #**	**Countries**
1	Swaziland Tajikistan		2	India
3	Korea, Rep.	Sweden	4	Brazil
	Germany	Switzerland		Argentina
	Australia	United Kingdom		
	Austria	France		
	Belgium	Spain		
	Canada	Hong Kong SAR, China		

	Denmark Netherlands Norway Finland	Singapore Iceland Ireland Israel	
5	China		6	United States Japan
7	Bolivia Brunei Darussalam Indonesia Lebanon United Arab Emirates Nigeria Sri Lanka Zambia		8	Bosnia and Herzegovina
9	Italy Russian Federation Turkey Saudi Arabia Qatar El Salvador Kuwait Oman Bahrain Iran, Islamic Rep. Jordan Malaysia Egypt, Arab Rep. Algeria	Portugal Poland Romania Bulgaria Hungary Estonia Croatia Serbia Montenegro Macedonia, FYR Ukraine Moldova Slovenia Czech Republic	Armenia Azerbaijan Kazakhstan Kyrgyz Republic South Africa Botswana Ghana Gabon Barbados Trinidad and Tobago Namibia Chile Uruguay Colombia	

	Morocco	Slovak Republic	Costa Rica
	Tunisia	Georgia	Ecuador
	Brunei Darussalam	Latvia	Dominican Republic
	Philippines	Lithuania	Panama
	Thailand	Greece	Peru
	Vietnam	Cyprus	Puerto Rico
	Mexico	Albania	
10	Luxemburg		

5.5.2.3 Approach 3: Applying CityBlock with complete linkage clustering

In this section the CityBlock with complete linkage clustering will be applied into the large size real-world sample and the formed clusters are identified as shown below.

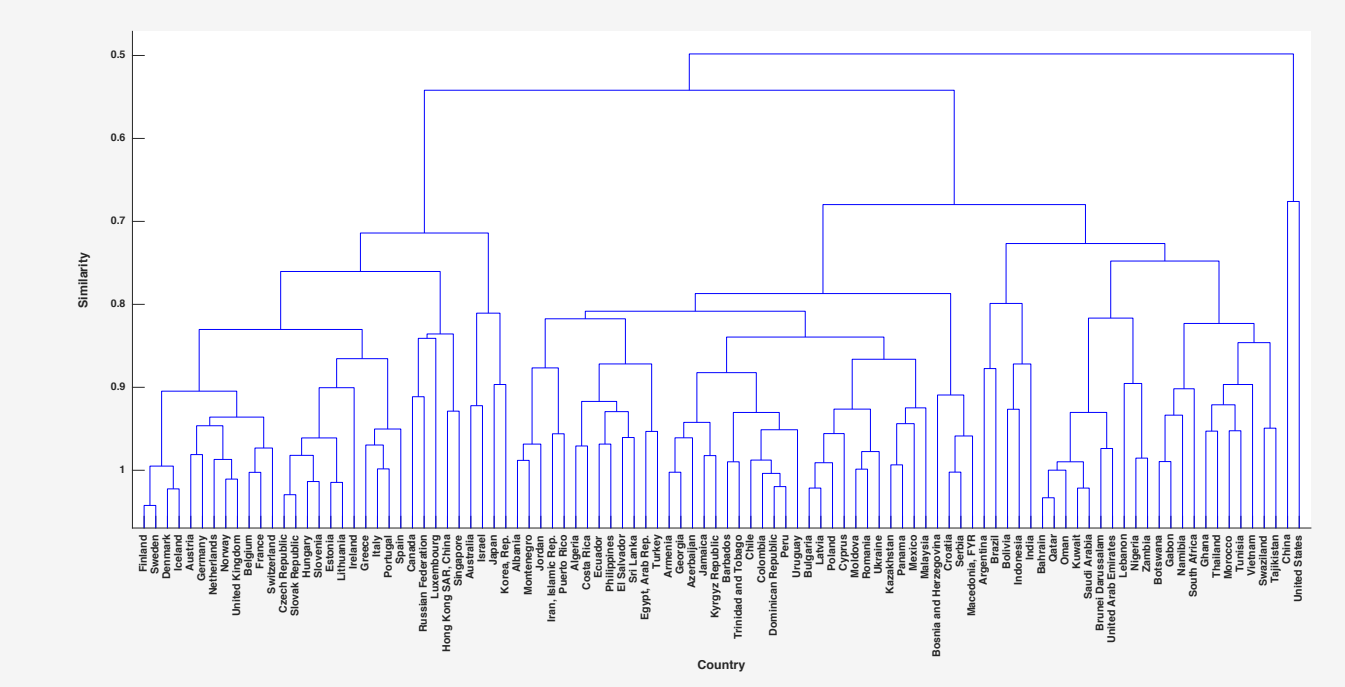

Figure (5.14) Dendrogram of clustering the large size real-world sample countries using CityBlock with complete linkage clustering

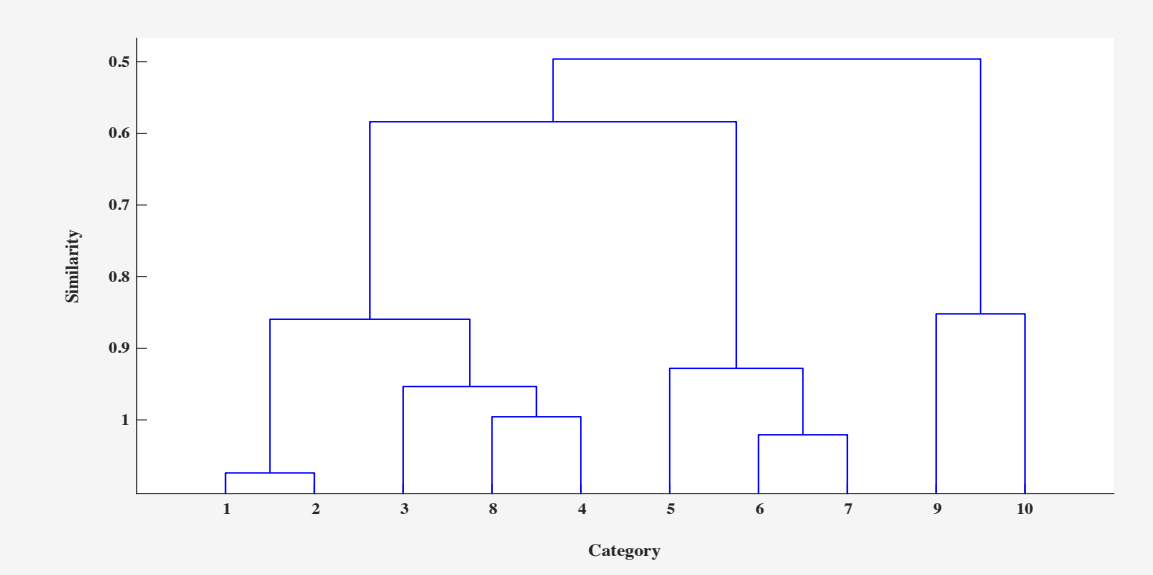

Figure (5.15) Dendrogram of clustering the large size real-world sample countries using CityBlock with complete linkage clustering in ten categories

Table (5.35) Assigning the countries to clusters using CityBlock with complete linkage clustering in ten categories

Cluster #	Countries	Cluster #	Countries
1	United States	2	Canada Hong Kong SAR, China Luxembourg Russian Federation Singapore
3	Australia Israel Japan Korea, Rep.	4	Brazil India Indonesia Argentina Bolivia

5	Austria	Ireland	6	Bahrain
	Belgium	Italy		Brunei Darussalam
	Czech Republic	Lithuania		Kuwait
	Denmark	Netherlands		Lebanon
	Estonia	Norway		Nigeria
	Finland	Portugal		Oman
	France	Slovak Republic		Qatar
	Germany	Slovenia		Saudi Arabia
	Greece	Spain		United Arab Emirates
	Hungary	Sweden		Zambia
	Iceland	Switzerland		
		United Kingdom		
7	Bosnia and Herzegovina		8	Botswana
	Croatia			Gabon
	Macedonia, FYR			Ghana
	Serbia			Morocco
				Namibia
				South Africa
				Swaziland
				Tajikistan
				Thailand
				Tunisia
				Vietnam
9	Albania	Egypt, Arab Rep.		Montenegro
	Algeria	El Salvador		Panama
	Armenia	Georgia		Peru
	Azerbaijan	Iran, Islamic Rep.		Philippines
	Barbados	Jamaica		Poland
	Bulgaria	Jordan		Puerto Rico
	Chile	Kazakhstan		Romania

	Colombia	Kyrgyz Republic	Sri Lanka
	Costa Rica	Latvia	Trinidad and Tobago
	Cyprus	Malaysia	Turkey
	Dominican Republic	Mexico	Ukraine
	Ecuador	Moldova	Uruguay
10	China		

5.5.2.4 Approach 4: Applying CityBlock with average linkage clustering

The last approach that will be applied into the large size real-world sample is the CityBlock coefficient with complete linkage clustering and the following results are obtained.

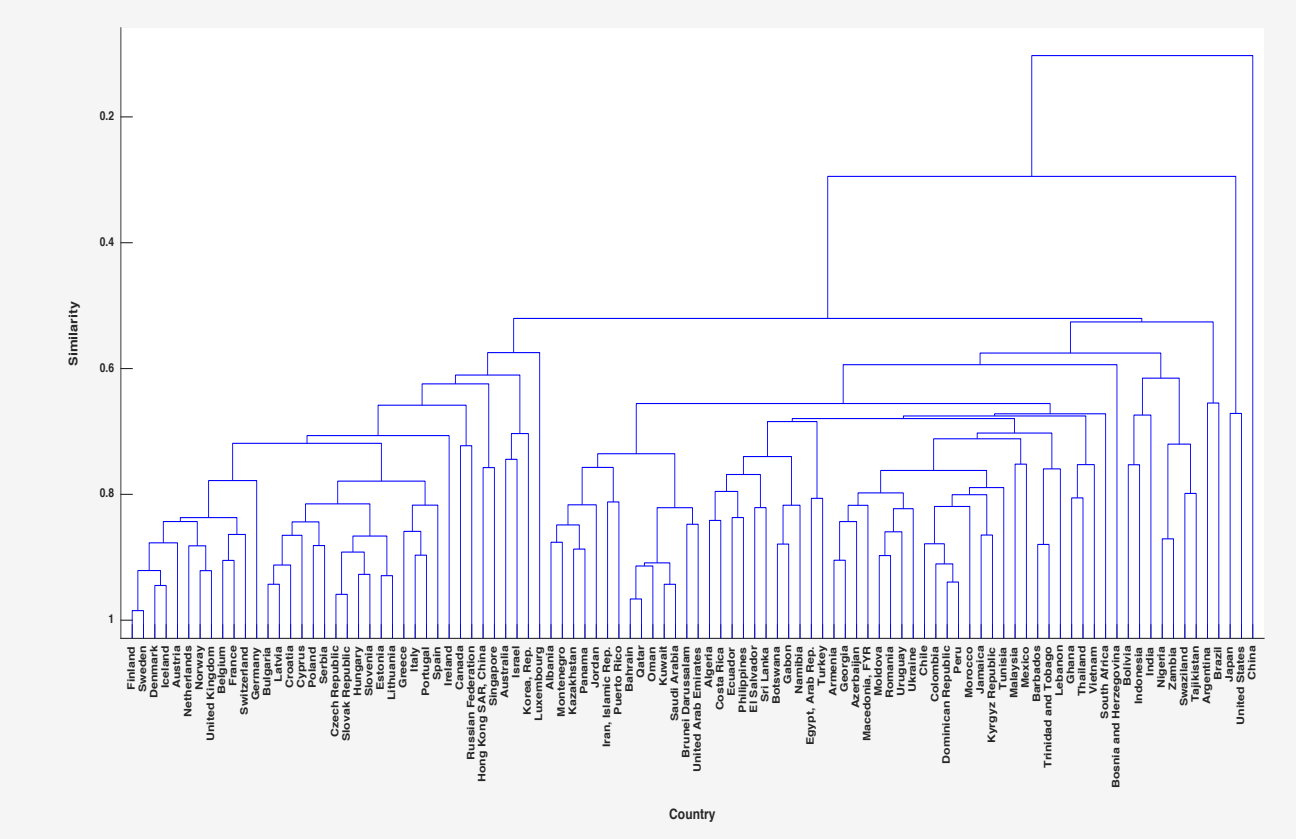

Figure (5.16) Dendrogram of clustering the large size real-world sample countries using CityBlock with complete linkage clustering

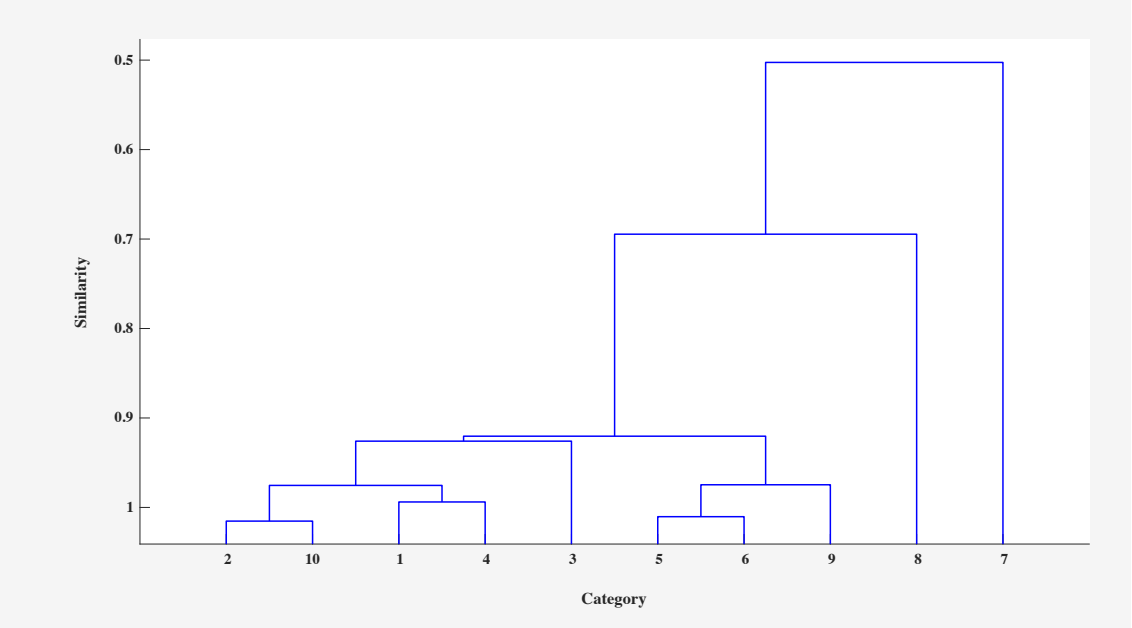

Figure (5.17) Dendrogram of clustering the large size real-world sample countries using CityBlock with average linkage clustering in ten categories

Table (5.36) Assigning the countries to clusters using CityBlock with average linkage clustering in ten categories

Cluster #	**Countries**	**Cluster #**	**Countries**
1	Nigeria Swaziland Tajikistan Zambia	2	China
3	Australia Israel Korea, Rep.	4	Argentina Brazil
5	Austria Belgium	Germany Greece	Poland Portugal

	Bulgaria	Hong Kong SAR, China	Russian Federation
	Canada	Hungary	Serbia
	Croatia	Iceland	Singapore
	Cyprus	Ireland	Slovak Republic
	Czech Republic	Italy	Slovenia
	Denmark	Latvia	Spain
	Estonia	Lithuania	Sweden
	Finland	Netherlands	Switzerland
	France	Norway	United Kingdom
6	United States Japan	7	Bolivia India Indonesia
8	Bosnia and Herzegovina		
9	Albania	Ghana	Peru
	Algeria	Iran, Islamic Rep.	Philippines
	Armenia	Jamaica	Puerto Rico
	Azerbaijan	Jordan	Qatar
	Bahrain	Kazakhstan	Romania
	Barbados	Kuwait	Puerto Rico
	Botswana	Kyrgyz Republic	Saudi Arabia
	Brunei Darussalam	Lebanon	South Africa
	Chile	Macedonia, FYR	Sri Lanka
	Colombia	Malaysia	Thailand
	Costa Rica	Mexico	Trinidad and Tobago
	Dominican Republic	Moldova	Tunisia
	Ecuador	Montenegro	Turkey

	Egypt, Arab Rep.	Morocco	Ukraine
	El Salvador	Namibia	United Arab Emirates
	Gabon	Oman	Uruguay
	Georgia	Panama	Vietnam
10	Luxembourg		

Moreover, the results obtained from applying the different four approaches into the large size real-world sample can be summarized in the following table.

Table (5.37) Categorizing the countries to clusters based on four different clustering approaches

Country	Approach 1	Approach 2	Approach 3	Approach 4
Korea, Rep.	3	3	3	3
United States	1	6	1	6
Japan	1	6	3	6
Israel	3	3	3	3
Germany	5	3	5	5
Denmark	5	3	5	5
Sweden	5	3	5	5
China	2	5	10	2
Finland	5	3	5	5
Switzerland	5	3	5	5
Netherlands	5	3	5	5
Iceland	5	3	5	5
Norway	5	3	5	5
Austria	5	3	5	5
France	5	3	5	5
United Kingdom	5	3	5	5

Belgium	5	3	5	5
Singapore	5	3	2	5
Czech Republic	6	9	5	5
Slovenia	6	9	5	5
Australia	3	3	3	3
Luxembourg	5	10	2	10
Canada	5	3	2	5
Estonia	6	9	5	5
Ireland	5	3	5	5
Hungary	6	9	5	5
Slovak Republic	6	9	5	5
Spain	6	3	5	5
Argentina	4	4	4	4
Italy	6	9	5	5
Russian Federation	6	9	2	5
Portugal	6	9	5	5
Hong Kong SAR, China	5	3	2	5
Malaysia	10	9	9	9
Lithuania	6	9	5	5
Greece	6	9	5	5
Brazil	4	4	4	4
Poland	8	9	9	5
Barbados	10	9	9	9
Uruguay	10	9	9	9
Turkey	8	9	9	9
Ukraine	8	9	9	9
Serbia	8	9	7	5
Latvia	8	9	9	5
Croatia	8	9	7	5

Lebanon	7	7	6	9
Bulgaria	8	9	9	5
Azerbaijan	9	9	9	9
Chile	10	9	9	9
Costa Rica	9	9	9	9
Bosnia and Herzegovina	8	8	7	8
Trinidad and Tobago	9	9	9	9
Romania	8	9	9	9
Colombia	10	9	9	9
Morocco	10	9	8	9
Moldova	8	9	9	9
Puerto Rico	9	9	9	9
Macedonia, FYR	8	9	7	9
Montenegro	9	9	9	9
Cyprus	8	9	9	5
Dominican Republic	10	9	9	9
Tunisia	10	9	8	9
Peru	10	9	9	9
Armenia	4	9	9	9
Thailand	10	9	8	9
Mexico	9	9	9	9
Georgia	8	9	9	9
South Africa	10	9	8	9
Jamaica	10	9	9	9
Bahrain	9	9	6	9
Qatar	9	9	6	9
Vietnam	10	9	8	9
Kazakhstan	9	9	9	9
Jordan	8	9	9	9

India	7	2	4	7
United Arab Emirates	9	7	6	9
Egypt, Arab Rep.	8	9	9	9
Bolivia	7	7	4	7
Ecuador	9	9	9	9
Albania	9	9	9	9
Indonesia	7	7	4	7
Philippines	9	9	9	9
Tajikistan	10	1	8	1
Botswana	9	9	8	9
Kyrgyz Republic	8	9	9	9
Oman	9	9	6	9
Iran, Islamic Rep.	10	9	9	9
Sri Lanka	7	7	9	9
El Salvador	9	9	9	9
Kuwait	9	9	6	9
Namibia	10	9	8	9
Brunei Darussalam	9	7	6	9
Panama	9	9	9	9
Saudi Arabia	9	9	6	9
Algeria	9	9	9	9
Nigeria	7	7	6	1
Ghana	10	9	8	9
Swaziland	10	1	8	1
Zambia	7	7	6	1
Gabon	9	9	8	9

The results from Table (5.36) indicate that applying different clustering approaches does not have big effects on categorizing the countries - according to their entrepreneurial

attractiveness - into distinct clusters; i.e., the countries that are identified and categorized in one individual cluster using one clustering approach are similar, to a far extent, to those countries that are grouped into an individual cluster through a different clustering approach.

However, the adopted clustering algorithm in the developed model is the complete linkage clustering with the Euclidean distance similarity coefficient due to several reasons that were discussed in section 4.4.2 such as it uses the least similar pair factor to determine the inter-cluster similarity, the identified clusters are small and tightly bound, it prevents the merge of two clusters together for only high level of similarity, and like other similarity-based clustering algorithms, it is computer software-friendly.

CHAPTER SIX

Conclusions and Future Research

6.1 Conclusions

This research is proposing an algorithm to approach the facility location problem of entrepreneurial organizations with global orientations based on several similarity coefficient-based clustering models. In general, the developed model suggests that countries with similar attributes are classified and compiled together in distinctive groups. This process could assist the entrepreneurs/decision makers to construct a better viable decision to locate their facility within a flexible pool of potential countries that fit the scope and activities of the considered businesses. The final decision would rely on comprehensive decision-making attributes in, which ranking and favored locations also take place.

Classifying candidate countries based on a combination of location decision-making factors also reduces the influence of error in data collection and/or analysis in deciding a better potential location for the business. However, the set of decisive attributes has to be carefully composed in order not to exclude material factors. To do so, the most frequent considered location decision-making factors in the various available resources of data have to be extensively studied.

Determining the factors of location attraction to entrepreneurs is a crucial threshold in implementing the developed model both correctly and effectively. Inability to identify the most important factors would most likely yield misleading and false outcomes. On the other hand, in order to obtain reliable results, the essential decision-making factors that are tightly related to the considered entrepreneurial activity must be specified.

6.2 Future Research

Location decisions adopting the developed model in this research lead to identifying a group of potential countries to accommodate the new business. However, determining the best alternative within a single group of countries demands embedding additional decisive factors to decide between the alternatives among the group in accordance to the type and nature of the desired entrepreneurial activity.

More attention might be given to aligning the internal resources that exist within the start-up entrepreneurial firm with external business-attraction factors in the location decision-making process.

The process of the facility location decision-making for specialized entrepreneurial ventures (e.g., technological-based small companies) might be conducted in the same context by considering the specific decision-making factors that are related to the type of the business.

Another research scope could be applying the resulting classifications to help the regional development authorities in designing more attractive business sites for new entrepreneurial endeavors in more credible approaches.

REFERENCES

A.K. Jain, M.N. Murty and P.J. Flynn. 1999. "Data Clustering: A Review". ACM Computing Surveys. 31(3).

Ajay Anandan. 2013. "Efficient Algorithms for Hierarchical Agglomerative Clustering". University of Alberta.

Alain Verbeke, Liena Kano. 2012. "An internalization theory rationale for MNE regional strategy". Multinational Business Review. 20(2): 135-152.

Alex Rialp-Criado, Inmaculada Galvan-Sanchez and Sonia M. Sua´rez-Ortega. 2010. "A configuration-holistic approach to born-global firms strategy formation process". European Management Journal. 28: 108– 123.

Ali Aslan Gumusay. 2015. "Entrepreneurship from an Islamic Perspective". Journal of Business Ethics. 130(1): 199-208.

Alvaro Cuervo-Cazurra. 2011. "Selecting the country in which to start internationalization: The non-sequential internationalization model". Journal of World Business. 46: 426–437.

Andreea N. Kiss, Wade M. Danis and S. Tamer Cavusgil. 2012. "International entrepreneurship research in emerging economies: A critical review and research agenda". Journal of Business Venturing. 27: 266–290.

Anil K. Jain, Richard C. Dubes. 1988. "Algorithms for Clustering Data".

Anni Tuppura, Sami Saarenketo, Kaisu Puumalainen, Ari Jantunen and Kalevi Kylaheiko. 2008. "Linking knowledge, entry timing and internationalization strategy". International Business Review. 17: 473–487.

Azizallah Roudini, Mohd. Hassan Mohd. Osman. 2012. "The Role of International Entrepreneurship Capability on International Performance in Born Global Firms". iBusiness. 4: 126-135.

Benjamin M. Oviatt, Patricia Phillips McDougall and Marvin Loper. 1995. "Global Start-Ups: Entrepreneurs on a Worldwide Stage". The Academy of Management Executive (1993-2005) 9.2, Careers in the 21st Century: 30-44.

Birgit Hagen, Antonella Zucchella. 2014. "Born Global or Born to Run? The Long-Term Growth of Born Global Firms". Management International Review. 54: 497-525.

Breda McCarthy 2003. "Strategy is personality-driven, strategy is crisis-driven: insights from entrepreneurial firms". Management Decision. 41(4): 327 – 339.

C. Gorter. 2000. “Migrant Entrepreneurs in East Indonesia: A Schumpeterian Perspective“. Tinbergen Institute Discussion Paper 00-082/3.

C.L. Yang, S.P. Chuang, R.H. Huang and C.C. Tai. 2008. “Location Selection Based on AHP/ANP Approach”. IEEM 2008: the IEEE International Conference on Industrial Engineering and Engineering Management: 8 to 11 December, 2008, Singapore: 1148-1153.

Cathy H. Yang. 2004. “Identifying and Testing the Decision Making Factors Related to ‘Key Industries’ Choice of Location”.

Chiung-Hui Tseng, Patriya Tansuhaj, William Hallagan and James McCullough. 2007. “Effects of firm resources on growth in multinationality”. Journal of International Business Studies. 38: 961–974.

Christian Kingombe, Milford Bateman and Dirk Willem te Velde. 2010. “Review of the most recent literature on Entrepreneurship and SMEs”.

Christiane Gebhardt, Markus C. Pohlmann. 2013. “Managing the organisation 2.0: Entrepreneurial spirit and general management competences as early indicators for cluster success and sustainable regional development: Findings from the German Entrepreneurial Regions Programme”. Journal of High Technology Management Research. 24: 153–160.

Christos Kolympiris, Nicholas Kalaitzandonakes and Douglas Miller. 2015. “Location choice of academic entrepreneurs: Evidence from the US biotechnology industry”. Journal of Business Venturing. 30: 227–254.

Colm O’Gorman, Lorraine McTiernan. 2000. “Factors Influencing the Internationalization Choices of Small and Medium-Sized Enterprises: The Case of the Irish Hotel Industry”. Enterprise & Innovation Management Studies. 1(2): 141-151.

Craig C. Julian, Zafar U. Ahmed. 2012. “Factors impacting international entrepreneurship in Malaysia”. Journal of Small Business and Enterprise Development. 19(2): 229-245.

Craig S. Galbraith, Carlos L. Rodriguez and Alex F. DeNoble. 2008. “SME Competitive Strategy and Location Behavior: An Exploratory Study of High-Technology Manufacturing”. Journal of Small Business Management. 46(2): 183–202.

Daniela Grieco. 2012. “Economics Perspectives on the Entrepreneurial Decision”. International Journal of the Economics of Business. 19(3): 451–467.

Dara Szyliowicz, Tiffany Galvin. 2010. “Applying broader strokes: Extending institutional perspectives and agendas for international entrepreneurship research”. International Business Review. 19: 317–332.

David A Schilling, Vaidyanathan Jayaraman and Reza Barkhi. 1993. “A review of covering problems in facility location”. Location Services. 1(1): 25-55.

Dean A. Shepherd, Trenton A. Williams and Holger Patzelt. 2015. “Thinking About Entrepreneurial Decision Making: Review and Research Agenda”. Journal of Management. 41(1): 11–46.

Dengsheng Zhang and Guojun Lu. 2003. “Evaluation of similarity measurement for image retrieval”. IEEE International Conference: Neural Networks & Signal Processing, Nanjing, China, December 14-17, 2003.

Domingo Ribeiro Soriano and Kun-Huang Huarng. 2013. “Innovation and entrepreneurship in knowledge industries”. Journal of Business Research. 66(10): 1964-1969.

Federico Pablo-Martí, Antonio García-Tabuenca, Juan Luis Santos, María Teresa Gallo and Tomás Mancha. 2012. “Entrepreneurship and Firms’ Mobility”.

Fernando Angel Garza. 2013. “International Business Locations: A Frmework and Model for Entrepreneurial “Reshoring””.

Francesco Quatraro, Marco Vivarelli. 2013. “Drivers of Entrepreneurship and Post-Entry Performance of Newborn Firms in Developing Countries”. Discussion Paper 7436.

G. Jayakrishnan Nair, T.T. Narendran. 1998. “Case: A clustering algorithm for cell formation with sequence data”. International Journal of Production Research. 36(1): 157-180.

Gabriele Suder, Peter W. Liesch, Satoshi Inomata, Irina Mihailova and Bo Meng. 2015. “The evolving geography of production hubs and regional value chains across East Asia: Trade in value-added”. Journal of World Business. 50: 404–416.

Gina Beim, Moren Levesque. 2003. “Country Selection for New Business Venturing: A Multiple Criteria Decision Analysis”. Technical Memorandum Number 776R. Department of Operations, Weatherhead School of Management, Case Western Reserve University.

Giusy Cannone, Gaia Costantino, Alessia Pisoni and Alberto Onetti. 1989. “Drivers of international development for born global companies founded by Italian entrepreneurs”.

H Charles Romesburg. 2012. “Cluster Analysis for Researchers”.

Hamid K. Seifoddini. 1976. “Single linkage versus average linkage clustering in machine cells formation applications”. Journal Computers and Industrial Engineering. 16(3): 419-426.

Hanadi Salameh. 2001. “Clustering Relational Databases”. Bell & Howell Information and Learning Company.

Ian Durbach, Risto Lahdelma and Pekka Salminen, 2014. “The analytic hierarchy process with stochastic judgments”. European Journal of Operational Research. 238: 552-559.

James W. Carland, Frank Hoy, William R. Boulton and Jo Ann C. Carland 1984. “Differentiating Entrepreneurs from Small Business Owners: A Conceptualization”. The Academy of Management Review. 9(2): 354-359.

Jerry T. Parwada. 2008. “The Genesis of Home Bias? The Location and Portfolio Choices of Investment Company Start-Ups”. Journal of Financial and Quantitative Analysis. 43(1): 245–266.

Jiaqin Yang and Huei Lee. 1997. “An AHP decision model for facility location selection. Facilities”. 15(9): 241.

Joanne Freeman, Chris Styles and Meredith Lawley. 2012. “Does firm location make a difference to the export performance of SMEs?”. International Marketing Review. 29(1): 88-113.

John D. Mittelstaedt, William A. Ward and Edward Nowlin. 2006. “Location, industrial concentration and the propensity of small US firms to export”. International Marketing Review. 23(5): 486 – 503.

John H. Dunning. 2009. “Location and the Multinational Enterprise: A Neglected Factor?”. Journal of International Business Studies. 40(1): 5-19.

Jolita Sinkiene, Saulius Kromalcas. 2010. “Concept, Directions and Practice of City Attractiveness Improvement”. Public Policy and Administration. 31: 147-154.

Juan Carlos Dıaz Casero, Manuel Almodovar Gonzalez, Marıa de la Cruz Sanchez Escobedo, Alicia Coduras Martınez and Ricardo Hernandez Mogollon. 2013. “Institutional variables, entrepreneurial activity and economic development”. Management Decision. 51(2): 281-305.

Jun Wang, Catherine Roze. 1994. “Formation of Machine Cells and Part Families in Cellular Manufacturing: A Linear Integer Programming Approach”. Industrial Technology, Proceedings of the IEEE International Conference.

Katharina Pijnenburg, Konstantin A. Kholodilin. 2014. “Do Regions with Entrepreneurial Neighbours Perform Better? A Spatial Econometric Approach for German Regions”. Regional Studies. 48(5): 866-882.

Keith J. Perks, Mathew Hughes. 2008. “Entrepreneurial decision-making in internationalization: Propositions from mid-size firms”. International Business Review. 17: 310-330.

Khan Md. Saiful Islam and Bhaba R. Sarker. 2000. “A similarity coefficient measure and machine-parts grouping in cellular manufacturing systems”. International Journal of Production Research. 38(3): 699-720.

Leora Klapper, Raphael Amit, Mauro F. Guillén and Juan Manuel Quesada. 2007. "Entrepreneurship and Firm Formation Across Countries". The World Bank, Policy Research Working Paper 4313.

Linzy D. Albert, James H. Kellow. 1969. "Decision-Makers' Reactions to Plant Location Factors: An Appraisal". Land Economics. 45(3): 376-381.

M. P. Chandrasekharan and R. Rajagopalan. 1989. "GROUPABIL1TY: an analysis of the properties of binary data matrices for group technology". International Journal of Production Research. 27(6): 1035-1052.

M.Şebnem Ensaria, Melisa Erdilek Karabayb. 2014. "10th International Strategic Management Conference What Helps to Make SMEs Successful in Global Markets?". Procedia - Social and Behavioral Sciences. 150: 192 – 201.

Magdalena Zwolińska-Ligaj. 2015. "The Influence of Natural Valuable Areas on the Development of Entrepreneurship (Based on the Example of Lublin Voivodeship)". International Conference, Economic Science for Rural Development 39, Jelgava, LLU ESAF: 23-24, 34-46.

Mahmood Nooraie. 2008. "Decision magnitude of impact and strategic decision-making process output". Management Decision. 46(4): 640 – 655.

Manuel Espitia-Escuer, Lucía I. García-Cebrián and Antonio Muñoz-Porcar. 2015 "Location as a competitive advantage for entrepreneurship an empirical application in the Region of Aragon (Spain)". International Entrepreneurship Management Journal. 11: 133–148.

Marian V. Jones, Nicole Coviello and Yee Kwan Tang. 2011. "International Entrepreneurship research (1989–2009): A domain ontology and thematic analysis". Journal of Business Venturing. 26: 632–659.

Marja Roza, Frans A.J. Van den Bosch and Henk W. Volberda. 2011. "Offshoring strategy: Motives, functions, locations, and governance modes of small, medium-sized and large firms". International Business Review. 20: 314–323.

Martin Andersson, Börje Johansson. 2007. "Product Innovation, Export Entrepreneurship and Regional Characteristics - an analysis of innovation ideas in regions".

Maryann Feldman, Johanna Francis and Janet Bercovitz. 2005. "Creating a Cluster While Building a Firm: Entrepreneurs and the Formation of Industrial Clusters". Regional Studies. 39(1): 129-141.

Maury D. Granger, Glenn C. Blomquist. 1999. "Evaluating the influence of amenities on the location of manufacturing establishments in Urban areas". Urban Studies. 36(11): 1859-1873.

Mercedes Delgado, Michael E. Porter and Scott Stern. 2010. “Clusters and entrepreneurship”. Journal of Economic Geography. 10: 495–518.

Michael S. Dah, Olav Sorenson. 2012. “Home Sweet Home: Entrepreneurs' Location Choices and the Performance of Their Ventures”. Management Science. 58(6): 1059-1071.

Michael W-P Fortunato, Theodore R. Alter. 2009. “Against the Grain: What Motivates Entrepreneurs to Locate in Pennsylvania’s Non-Metropolitan Cities and Boroughs”.

Ming-Jer Chen, Donald C. Hambrick. 1995. “Speed, Stealth, and Selective Attack: How Small Firms Differ from Large Firms in Competitive Behavior”. The Academy of Management Journal. 38(2): 453-482.

Mohd Sobri Minai, Esuh Ossai-Igwe Lucky 2011. “The Conceptual Framework of the Effect of Location on Performance of Small Firms”. Asian Social Science. 7(12).

Nicole E. Coviello, Patricia P. McDougall and Benjamin M. Oviatt. 2011. “The emergence, advance and future of international entrepreneurship research- An introduction to the special forum”. Journal of Business Venturing. 26: 625–631.

Niina Nummela, Sami Saarenketo, Paivi Jokela and Sharon Loane. 2014. “Strategic Decision-Making of a Born Global: A Comparative Study From Three Small Open Economies”. Management International Review. 54: 527–550.

Odilia Yim, Kylee T. Ramdeen. 2015. “Hierarchical Cluster Analysis: Comparison of Three Linkage Measures and Application to Psychological Data”. The Quantitative Methods for Psychology. 11(1).

Oliver Falck, Michael Fritsch and Stephan Heblich. 2008. “The Apple Doesn´t Fall Far From the Tree: Location of Start-Ups Relative to Incumbents”. Jena Economic Research Papers. 082.

Pascal Beckers, Robert C. Kloosterman. 2011. “Business spaces between entrepreneurs and opportunities. The impact of the local built environment and zoning regulations on businesses in Dutch cities”.

Pascal Huett, Matthias Baumb, Christian Schwens and Ruediger Kabst. 2014. “Foreign direct investment location choice of small- and medium-sized enterprises: The risk of value erosion of firm-specific resources”. International Business Review. 23: 952–965.

Paul Dettwilera, Peter Lindelöfb and Hans Löfsten. 2006. “Utility of location: A comparative survey between small new technology based firms located on and off Science Parks—Implications for facilities management”. Technovation. 26(4): 506–517.

Paul Westhead, Mike Wright and Deniz Ucbasaran. 2002. “International market selection strategies selected by ‘micro’ and ‘small’ firms”. Omega. 30(1): 51–68.

Pavlos Dimitratos, Irini Voudouris, Emmanuella Plakoyiannaki and George Nakos. 2012. "International entrepreneurial culture- Toward a comprehensive opportunity-based operationalization of international entrepreneurship". International Business Review. 21: 708–721.

Per Davidsson. 2005. "Researching Entrepreneurship".

Peter J Buckley, Timothy M Devinney and Jordan J Louviere. 2007. "Do managers behave the way theory suggests? A choice-theoretic examination of foreign direct investment location decision-making". Journal of International Business Studies. 38: 1069–1094.

Peter Nijkamp and Jos van Ommeren. 2004. "Drivers of entrepreneurial location as an innovative act".

Qinqin Zheng, Yadong Luo and Vladislav Maksimov. 2015. "Achieving legitimacy through corporate social responsibility: The case of emerging economy firms". Journal of World Business. 50: 389–403.

R.Tamilselvi, B.Sivasakthi and R.Kavitha. 2015. "A comparison of various clustering methods and algorithms in data mining". International Journal of Multidisciplinary Research and Development. 2(5): 32-36.

Raphael Suire, Jerome Vicenteb. 2014. "Clusters for life or life cycles of clusters: in search of the critical factors of clusters' resilience". Entrepreneurship & Regional Development. 26: 1–2, 142–164.

Robert McIntyre. 2001. "The Role of Small and Medium Enterprises in Transition: Growth and Entrepreneurship".

Roger K. Blashfield, 1976. "Mixture Model Tests of Cluster Analysis: Accuracy of Four Agglomerative Hierarchical Methods". Psychological Bulletin. 83(3): 377-388.

Roy Suddaby, Garry D. Bruton and Steven X. Si. 2015. "Entrepreneurship through a qualitative lens: Insights on the construction and/or discovery of entrepreneurial opportunity". Journal of Business Venturing. 30: 1–10.

Sami Mahroum. 2000. "Highly skilled globetrotters: mapping the international migration of human capital".

Sanjeev Agarwal, Sridhar N. Ramaswami. 1992. "Choice of Foreign Market Entry Mode: Impact of Ownership, Location and Internalization Factors". Journal of International Business Studies. 23.1: 1-27.

Seyyed Amin Terouhid, Robert Ries and Maryam Mirhadi Fard. 2012. "Towards Sustainable Facility Location – A Literature Review". Journal of Sustainable Development. 5(7).

Shelley M. Kimelberg, Elizabeth Williams. 2013. “Evaluating the Importance of Business Location Factors: The Influence of Facility Type”. Growth and Change. 44(1): 92-117.

Skokan Karel, Pawliczek Adam and Piszczur Radomír. 2013. “Strategic Planning and Business Performance of Micro, Small and Medium-Sized Enterprises”. Journal of Competitiveness. 5(4): 57-72.

Srinath Dissanayake. 2013. “Entrepreneurship and the Entrepreneur: A Synthesis View”. Munich Personal RePEc Archive. Paper No. 49255.

Stephan Gerschewski, Elizabeth L. Rose and Valerie J. Lindsay. 2015. “Understanding the drivers of international performance for born global firms: An integrated perspective”. Journal of World Business. 50: 558–575.

Susan Hesse Owen, Mark S. Daskin. 1998. “Strategic facility location: A review”. European Journal of Operational Research. 111: 423-447.

Tarun Gupta, Hamid Seifoddini. 1990. “Production data based similarity coefficient for machine-component grouping decisions in the design of a cellular manufacturing systems”. International Journal of Production Resources. 28(7): 1247-1962.

Tatiana Manolova, Candida Brush and Linda Edelman. 2011. “Initial resource assembly in new ventures: Does location matter?”. Psicothema. 23(3): 439-445.

Tian Zhang. 1997. “Data Clustering for very large database plus applications”.

Tim Mazzarol Stephen Choo. 2003. “A study of the factors influencing the operating location decisions of small firms”. Property Management. 21(2): 190 – 208.

Tim Rogmans. 2013. “Location and operation mode decision making in the Middle East: a case study approach”. Journal of Strategy and Management. 6(2): 190-206.

Virginia Carlson. 2000. “Studying Firm Locations: Survey Responses vs. Econometric Models”. The Journal of Regional Analysis & Policy. 30: 1-22.

Wim Naude. 2010. “Entrepreneurship, developing countries, and development economics: new approaches and insights”. Small Business Economics. 34: 1–12.

Wojciech Jarczewski. 2008. “Pro-investment Activities of Local Governments and Location Decisions of Entrepreneurs”. EUROPA XXI 17: 183-190.

Yong Yin, Kazuhiko Yasuda. 2006. “Similarity coefficient methods applied to the cell formation problem: A taxonomy and review”. International Journal of Production Economics. 101: 329–352.

APPENDICES

APPENDIX A:

Codes of MATLAB for Real-world Example

1. Code to obtain the dendrogram of the developed model for the real-world example using Euclidean distance with complete linkage clustering

```
- filename='G20';
- B=xlsread(filename);
- format long;
- for i=1:27
- B(:,i)=(B(:,i)-min(B(:,i)))./(max(B(:,i))-min(B(:,i)));
- end
- normalization
- D=pdist(B)
- tree=linkage(D,'Complete');
- [~, T]=dendrogram(tree,20)
```

2. Code to obtain the dendrogram of the of clustering the real-world example countries using Euclidean distance with complete linkage clustering in six categories

```
- filename='G20';
- B=xlsread(filename);
- format long;
- for i=1:27
- B(:,i)=(B(:,i)-min(B(:,i)))./(max(B(:,i))-min(B(:,i)));
- end
- normalization
- D=pdist(B)
- tree=linkage(D,'Complete');
- [~, T]=dendrogram(tree,6)
```

APPENDIX B:

Codes of MATLAB for the Effect of Number of Location Decision-making Factors Using the Real-world Example

1. Code to obtain the dendrogram of the G20 countries clustering based on three decision-making factors

```
- filename='G20-3 Factors';
- B=xlsread(filename);
- format long;
- for i=1:3
- B(:,i)=(B(:,i)-min(B(:,i)))./(max(B(:,i))-min(B(:,i)));
- end
- normalization
- D=pdist(B)
- tree=linkage(D,'Complete');
- [~, T]=dendrogram(tree,20)
```

2. Code to obtain the dendrogram of the G20 countries clustering based on six decision-making factors

```
- filename='G20-6 Factors';
- B=xlsread(filename);
- format long;
- for i=1:6
- B(:,i)=(B(:,i)-min(B(:,i)))./(max(B(:,i))-min(B(:,i)));
- end
- normalization
- D=pdist(B)
- tree=linkage(D,'Complete');
- [~, T]=dendrogram(tree,20)
```

APPENDIX C:

Codes of MATLAB for the Large Size Real-world Sample

1. Code to obtain the dendrogram of the developed model for the large size real-world sample using Euclidean distance with complete linkage clustering

```
- filename='100-Countries';
- B=xlsread(filename);
- format long;
- for i=1:27
- B(:,i)=(B(:,i)-min(B(:,i)))./(max(B(:,i))-min(B(:,i)));
- end
- normalization
- D=pdist(B)
- tree=linkage(D,'Complete');
- [~, T]=dendrogram(tree,100)
```

2. Code to obtain the dendrogram of the of clustering the large size real-world sample using Euclidean distance with complete linkage clustering in ten categories

```
- filename='100-Countries';
- B=xlsread(filename);
- format long;
- for i=1:27
- B(:,i)=(B(:,i)-min(B(:,i)))./(max(B(:,i))-min(B(:,i)));
- end
- normalization
- D=pdist(B)
- tree=linkage(D,'Complete');
- [~, T]=dendrogram(tree,10)
```

APPENDIX D:

Codes of MATLAB for Applying Other Clustering Approaches into the Large Size Real-world Sample

1. Code to obtain the dendrogram for the large size real-world sample using Euclidean distance with average linkage clustering

```
- filename='100-Countries';
- B=xlsread(filename);
- format long;
- for i=1:27
- B(:,i)=(B(:,i)-min(B(:,i)))./(max(B(:,i))-min(B(:,i)));
- end
- normalization
- D=pdist(B)
- tree=linkage(D,'Average');
- [~, T]=dendrogram(tree,100)
```

2. Code to obtain the dendrogram for the large size real-world sample using CityBlock distance with complete linkage clustering

```
- filename='100-Countries';
- B=xlsread(filename);
- format long;
- for i=1:27
- B(:,i)=(B(:,i)-min(B(:,i)))./(max(B(:,i))-min(B(:,i)));
- end
- normalization
- D=pdist(B, cityblock)
- tree=linkage(D,'Complete');
- [~, T]=dendrogram(tree,100)
```

3. Code to obtain the dendrogram for the large size real-world sample using CityBlock distance with average linkage clustering

```
- filename='100-Countries';
- B=xlsread(filename);
- format long;
- for i=1:27
- B(:,i)=(B(:,i)-min(B(:,i)))./(max(B(:,i))-min(B(:,i)));
- end
- normalization
- D=pdist(B, cityblock)
- tree=linkage(D,'Average');
- [~, T]=dendrogram(tree,100)
```

APPENDIX E:

Data for the top100 entrepreneurial countries (World Bank)

		Business Start-up Cost & Procedure						Financing SMEs			Tax Rates & Structure			Governmental Regulations & Policies
#	Country	Cost of business start-up procedures (% of GNI per capita)	Start-up procedures to register a business (number)	Time required to start a business (days)	Patent applications, nonresidents	Trademark applications, direct nonresident	Charges for the use of intellectual property, payments (current US$)	Firms using banks to finance investment (% of firms)	Lending interest rate (%)	Foreign direct investment, net (current US$)	Total tax rate (% of commercial profits)	Profit tax (% of commercial profits)	Taxes on goods and services (% value added of industry and services)	Exports of goods and services (% of GDP)
1	Albania	10.25	6.00	5.50	3.00	2,906.00	24,562,351.30	11.20	8.70	-969,143,530.00	36.50	14.10	0.00	27.10
2	Algeria	10.95	12.50	21.00	719.00	5,262.00	246,868,436.30	0.00	8.00	498,934,428.10	72.70	6.60	12.35	31.85
3	Argentina	12.40	14.00	25.00	4,173.00	14,352.00	1,909,402,685.10	30.30	24.90	-6,038,241,554.15	137.40	0.00	0.00	11.10
4	Armenia	1.00	2.00	3.00	2.00	3,485.00	0.00	17.40	17.60	-167,714,144.10	19.90	19.10	13.85	29.70
5	Australia	0.70	3.00	2.50	23,968.00	22,825.00	3,492,230,872.50	0.00	5.60	-40,946,217,336.90	47.60	26.00	6.35	19.80
6	Austria	0.30	8.00	22.00	271.00	3,018.00	1,512,195,649.50	0.00	0.00	9,030,757,481.30	51.70	16.80	13.45	53.40
7	Azerbaijan	2.15	2.50	4.00	0.00	4,765.00	23,301,500.00	27.10	17.50	-838,252,000.00	39.80	12.90	6.30	37.80
8	Bahrain	0.80	7.00	9.00	199.00	7,148.00	0.00	0.00	5.20	-486,835,106.40	13.50	0.00	0.10	72.40
9	Barbados	7.35	8.00	18.00	38.00	944.00	10,680,197.55	45.50	8.10	-251,528,354.15	34.70	19.50	15.75	36.90
10	Belgium	4.90	3.00	4.00	137.00	3,951.00	3,294,119,787.90	0.00	0.00	7,411,591,180.30	58.40	8.40	11.65	84.40
11	Bolivia	61.15	15.00	50.00	294.00	5,565.00	59,356,368.80	27.80	8.10	-911,104,145.45	83.70	0.00	0.00	43.70
12	Bosnia and Herzegovina	15.00	12.00	67.00	2.00	3,790.00	9,975,386.70	38.00	5.80	-227,079,582.90	23.30	7.20	22.40	33.75
13	Botswana	0.75	9.00	48.00	5.00	1,519.00	8,279,785.80	32.70	8.00	-309,118,589.80	25.10	21.50	5.65	49.70
14	Brazil	4.05	11.30	83.30	25,683.00	29,091.00	5,250,459,024.10	0.00	44.00	-61,576,211,015.00	69.20	0.00	7.65	13.00
15	Brunei Darussalam	5.85	12.50	59.00	91.00	1,074.00	0.00	0.00	5.50	171,289,904.90	8.70	0.80	0.00	73.60
16	Bulgaria	0.75	4.00	18.00	16.00	1,990.00	214,010,000.00	23.50	7.50	-1,364,380,000.00	27.00	5.00	16.60	66.50
17	Canada	0.40	2.00	3.50	31,283.00	29,680.00	9,384,234,274.10	0.00	2.80	25,435,858,098.60	21.10	3.90	2.70	31.50
18	Chile	0.70	7.00	5.50	2,653.00	8,963.00	1,544,972,045.10	44.80	5.50	-4,663,320,643.00	28.90	22.30	11.60	30.10
19	China	0.80	11.00	31.40	127,042.00	107,520.00	22,022,366,054.60	14.70	4.40	-62,058,344,182.60	67.80	0.00	7.75	22.40
20	Colombia	7.50	8.00	11.00	1,898.00	10,980.00	442,422,659.00	35.00	11.40	-7,724,264,828.30	69.70	22.50	7.40	14.70
21	Costa Rica	11.30	9.00	24.00	552.00	0	516,675,274.20	22.20	14.20	-2,541,924,546.40	58.00	19.30	9.35	35.30
22	Croatia	3.40	7.00	13.50	30.00	2,243.00	267,797,245.10	27.00	0.00	-175,731,670.20	20.00	0.00	19.15	49.40
23	Cyprus	12.40	6.00	8.00	0.00	1,267.00	0.00	0.00	0.00	5,103,147,988.60	24.40	9.30	14.45	53.10
24	Czech Republic	6.80	8.00	15.25	62.00	2,432.00	1,203,445,717.60	34.30	4.30	1,085,936,217.70	50.40	9.50	9.75	84.50
25	Denmark	0.20	4.00	4.25	206.00	1,785.00	1,220,774,891.20	0.00	0.00	9,222,194,658.80	24.50	18.70	17.40	53.30
26	Dominican Republic	16.65	7.00	14.50	245.00	4,114.00	88,900,000.00	39.10	14.90	-2,221,500,000.00	42.40	22.60	9.55	24.60
27	Ecuador	23.00	12.50	53.00	0.00	0.00	72,456,488.60	17.00	0.00	-1,060,063,337.30	33.00	16.10	0.00	21.10
28	Egypt, Arab Rep.	8.80	7.00	8.00	1,384.00	8161.00	284,500,000.00	8.90	11.60	-4,210,850,000.00	45.00	16.30	6.20	13.20
29	El Salvador	43.30	8.00	16.00	185.00	0.00	69,484,267.28	31.70	0.00	-369,937,551.85	38.80	20.15	10.96	25.91
30	Estonia	1.35	3.50	4.00	6.00	1560.00	42,458,387.60	32.70	4.50	76,269,960.20	49.40	8.40	1.00	79.80
31	Finland	1.05	3.00	14.00	126.00	1602.00	822,688,998.50	0.00	0.00	571,262,073.05	37.90	11.80	17.65	37.30
32	France	0.85	5.00	4.25	2,033.00	8210.00	13,914,984,295.80	0.00	0.00	-7,489,649,043.90	62.70	0.50	10.95	30.00
33	Gabon	14.00	7.00	50.00	0.00	0.00	0.00	0.00	0.00	0.00	45.70	20.90	0.00	45.90
34	Georgia	3.25	2.00	2.00	187.00	3911.00	6,985,876.40	22.00	12.50	-1,197,574,523.30	16.40	14.30	17.90	45.00
35	Germany	5.30	9.00	12.50	17,811.00	7543.00	8,916,969,815.90	0.00	0.00	62,608,628,930.50	48.80	23.20	7.65	46.90
36	Ghana	19.30	8.00	14.00	0.00	1366.00	0.00	21.20	0.00	-2,970,893,676.60	32.70	18.00	8.60	44.10
37	Greece	2.20	5.00	13.00	19.00	1349.00	290,625,614.50	0.00	0.00	672,556,082.00	49.60	19.70	16.35	30.10
38	Hong Kong SAR, China	1.30	2.50	2.00	12,350.00	24890.00	1,982,726,811.20	0.00	5.00	8,757,678,780.65	22.80	17.50	0.00	201.20
39	Hungary	7.80	4.00	5.00	73.00	1949.00	1,387,610,797.00	22.60	2.90	-387,951,609.20	48.40	11.80	19.60	88.65
40	Iceland	2.30	5.00	4.00	13.00	3048.00	97,138,938.20	0.00	7.60	-519,906,077.40	29.60	9.10	14.80	53.70
41	India	14.40	13.40	31.50	30,814.00	22098.00	4,376,310,359.60	30.30	10.00	0.00	60.60	0.00	3.80	24.10
42	Indonesia	20.50	13.00	50.15	7,321.00	11931.00	1,799,119,912.75	56.20	12.70	0.00	29.70	0.00	5.65	21.10
43	Iran, Islamic Rep.	2.90	8.00	15.00	119.00	3027.00	0.00	0.00	0.00	0.00	44.10	17.80	0.00	25.85
44	Ireland	0.25	4.00	6.00	58.00	1633.00	75,113,908,484.40	0.00	0.00	11,685,685,363.70	25.90	12.40	11.50	121.40
45	Israel	3.45	5.00	13.00	5,148.00	6982.00	1,061,600,000.00	46.40	3.40	-1,626,100,000.00	30.60	23.60	0.00	31.10
46	Italy	14.00	5.00	5.75	781.00	4468.00	4,321,299,926.30	0.00	4.10	7,410,936,028.00	64.80	19.50	10.30	30.20
47	Jamaica	4.95	2.00	9.00	122.00	1433.00	50,756,619.70	44.20	17.00	-790,088,374.60	35.20	14.80	11.60	30.10
48	Japan	7.50	8.00	10.20	60,030.00	24566.00	16,989,555,566.30	0.00	0.00	130,781,791,539.40	51.30	0.00	5.05	17.90
49	Jordan	21.05	7.00	12.00	339.00	4405.00	14,788,732.40	46.80	8.50	-1,273,802,816.90	29.50	13.10	0.00	37.80
50	Kazakhstan	0.30	5.00	7.50	271.00	7520.00	0.00	16.30	0.00	-6,935,985,001.75	29.20	16.20	0.00	28.60

		Governmental Regulations & Policies		Labor & Skills					Technology Advancement						Competition
#	Country	Trade in services (% of GDP)	Net official development assistance and official aid received (current US$)	Labor force with tertiary education (% of total)	Secondary education, vocational pupils	Government expenditure on education, total (% of government expenditure)	Wage and salaried workers, total (% of total employed)	Unemployment, total (% of total labor force)	High-technology exports (current US$)	Internet users (per 100 people)	Fixed broadband subscriptions (per 100 people)	Research and development expenditure (% of GDP)	Researchers in R&D (per million people)	Manufacturing, value added (current US$)	Listed domestic companies, total
1	Albania	34.20	274,725,000.00	14.80	24,288.00	0.00	37.95	16.05	15,123,662.00	63.30	7.60	0.00	0.00	574,512,070.40	0.00
2	Algeria	8.70	179,170,000.00	15.20	0.00	0.00	69.95	9.65	3,230,412.00	38.20	5.60	0.00	0.00	0.00	0.00
3	Argentina	5.50	40,845,000.00	21.05	0.00	15.40	76.15	7.65	1,930,301,086.50	69.40	16.10	0.60	1,198.00	84,192,294,420.90	93.00
4	Armenia	29.50	272,150,000.00	49.75	24,358.00	11.29	56.85	16.65	11,067,794.00	58.20	9.60	0.20	0.00	963,720,158.90	12.00
5	Australia	8.00	0.00	0.00	817,938.00	13.36	89.50	5.85	4,237,456,601.00	84.60	27.90	2.20	4,530.70	83,466,178,550.60	1,989.00
6	Austria	27.80	0.00	26.20	259,553.08	10.95	86.75	4.95	18,841,106,543.50	83.90	28.60	3.00	4,788.90	61,686,447,820.80	79.00
7	Azerbaijan	24.70	249,890,000.00	21.55	0.00	8.04	33.40	5.10	52,238,957.00	77.00	19.80	0.20	0.00	2,805,973,062.70	0.00
8	Bahrain	14.70	0.00	9.85	6,970.00	8.95	96.80	3.80	32,801,442.00	93.50	18.60	0.10	362.00	5,411,436,170.20	44.00
9	Barbados	45.35	16,200,000.00	0.00	0.00	14.62	83.20	11.80	26,483,703.00	76.10	27.20	0.00	0.00	169,875,000.00	20.00
10	Belgium	47.70	0.00	40.65	554,005.00	12.09	85.15	8.45	42,686,171,537.00	85.10	36.80	2.45	4,166.05	55,643,783,963.60	117.00
11	Bolivia	12.45	685,920,000.00	0.00	0.00	18.66	39.50	2.65	72,885,155.00	45.10	1.60	0.00	166.00	3,128,539,146.20	40.00
12	Bosnia and Herzegovina	13.30	577,075,000.00	14.70	122,273.00	0.00	73.60	27.65	96,042,566.00	65.10	16.60	0.30	241.70	1,796,027,218.60	767.00
13	Botswana	13.10	103,480,000.00	15.90	0.00	0.00	68.20	17.90	20,621,721.50	27.50	1.80	0.30	164.90	824,279,351.80	0.00
14	Brazil	5.90	1,028,105,000.00	13.05	874,889.00	15.42	68.65	6.65	8,848,309,553.00	59.10	12.20	1.20	698.10	173,262,496,618.50	345.00
15	Brunei Darussalam	14.80	0.00	0.00	6,861.00	9.86	0.00	3.80	54,307,486.50	71.20	8.00	0.00	0.00	1,994,444,923.15	0.00
16	Bulgaria	25.95	0.00	28.85	164,584.00	11.13	87.70	12.25	1,107,680,340.00	56.70	22.40	0.70	1,762.80	0.00	384.00
17	Canada	11.30	0.00	51.25	0.00	12.31	91.20	7.00	26,268,767,511.00	88.50	36.40	1.65	4,576.40	186,763,314,348.05	3,799.00
18	Chile	9.70	159,850,000.00	19.70	335,482.00	18.39	71.00	6.20	531,985,652.00	64.30	15.20	0.40	381.65	26,168,460,282.50	223.00
19	China	7.00	645,370,000.00	0.00	21,140,348.00	0.00	88.35	4.65	559,332,162,922.50	50.30	18.60	2.00	1,101.15	2,740,633,226,585.50	2,827.00
20	Colombia	6.30	1,037,700,000.00	25.40	348,734.00	16.19	46.90	9.85	729,911,512.50	55.90	11.20	0.25	150.00	32,611,190,369.40	69.00
21	Costa Rica	19.70	43,415,000.00	25.05	92,040.00	0.00	75.80	8.40	2,892,084,679.50	59.80	11.20	0.60	348.75	7,152,097,571.40	10.00
22	Croatia	33.20	0.00	23.45	0.00	8.85	83.15	17.00	760,238,175.00	69.80	23.20	0.80	1,482.90	6,010,877,506.10	186.00
23	Cyprus	74.60	0.00	42.55	0.00	15.61	82.25	15.70	26,759,530.00	71.70	22.40	0.50	760.75	1,074,282,513.80	84.00
24	Czech Republic	23.30	0.00	21.80	323,681.00	10.08	82.05	6.60	20,074,338,041.00	81.30	27.90	1.95	3,334.20	44,276,070,938.60	15.00
25	Denmark	39.00	0.00	32.30	134,687.00	15.07	91.05	6.80	9,318,204,877.00	96.30	42.50	3.10	7,183.20	37,621,210,630.10	0.00
26	Dominican Republic	15.90	157,030,000.00	21.50	0.00	20.65	53.30	15.00	180,074,209.00	51.90	6.40	0.00	0.00	9,430,599,455.20	0.00
27	Ecuador	5.50	153,765,000.00	21.20	593,331.00	10.66	50.95	4.40	97,006,477.00	48.90	9.20	0.35	160.80	14,802,545,000.00	0.00
28	Egypt, Arab Rep.	12.60	4,520,165,000.00	19.20	1,796,108.00	0.00	61.80	13.20	123,378,948.00	35.90	4.50	0.70	612.75	55,524,851,273.50	250.00
29	El Salvador	14.90	133,525,000.00	0.10	49.62	3.45	55.85	6.00	186,236,860.50	25.84	5.24	0.07	0.00	4,835,300,000.00	64.50
30	Estonia	43.30	0.00	39.30	15,539.00	13.55	90.90	8.15	1,040,763,519.00	88.40	28.70	1.55	3,311.45	2,968,592,658.30	0.00
31	Finland	21.80	0.00	39.90	252,191.00	12.37	86.30	8.40	3,632,887,080.00	92.70	31.70	3.25	7,086.90	32,500,831,762.20	0.00
32	France	19.50	0.00	35.80	1,109,601.00	9.80	88.65	10.15	132,183,573,785.00	84.70	41.30	2.25	4,185.45	242,860,153,044.30	490.00
33	Gabon	0.00	100,505,000.00	0.00	0.00	0.00	65.40	20.00	0.00	23.50	0.60	0.00	0.00	391,631,308.00	0.00
34	Georgia	34.70	605,030,000.00	30.70	21,350.00	7.99	38.40	14.00	34,733,393.00	45.20	14.60	0.10	573.50	1,546,732,472.60	134.00
35	Germany	16.80	0.00	27.75	1,338,507.38	10.82	88.90	5.15	184,283,164,631.00	87.60	37.20	2.85	4,390.15	681,526,006,432.30	555.00
36	Ghana	35.50	1,228,230,000.00	2.50	61,496.00	26.01	18.20	2.10	66,094,812.00	23.50	0.30	0.40	38.70	1,676,536,816.30	29.00
37	Greece	22.10	0.00	30.60	0.00	0.00	63.50	26.75	1,137,916,585.00	66.80	30.70	0.80	2,679.90	15,723,538,316.50	236.00
38	Hong Kong SAR, China	57.60	0.00	26.00	6,257.00	19.48	89.85	3.25	425,888,102.00	84.90	31.90	0.70	3,063.10	3,742,124,336.50	1,770.00
39	Hungary	31.30	0.00	25.10	133,788.00	9.52	88.70	9.00	11,759,507,186.00	72.80	27.40	1.40	2,586.70	26,497,004,278.85	45.00
40	Iceland	42.70	0.00	31.90	0.00	15.01	87.00	5.30	110,268,369.00	98.20	37.00	1.90	6,513.75	1,649,967,091.05	0.00
41	India	14.60	2,709,420,000.00	9.80	0.00	14.19	18.10	3.60	13,750,546,786.00	26.00	1.30	0.80	156.60	292,627,365,577.95	5,835.00
42	Indonesia	6.40	64,695,000.00	7.75	4,189,519.00	18.08	41.50	6.25	4,899,457,279.00	22.00	1.10	0.10	0.00	179,650,065,164.50	521.00
43	Iran, Islamic Rep.	0.00	104,970,000.00	0.00	0.00	19.35	55.65	12.85	618,441,975.50	44.10	10.90	0.30	714.80	54,264,954,285.50	318.00
44	Ireland	126.90	0.00	42.45	0.00	10.99	82.70	12.35	27,633,248,893.00	80.10	27.70	1.50	3,669.00	44,596,966,226.35	43.00
45	Israel	19.70	0.00	0.00	146,076.00	13.38	87.35	6.20	9,938,048,000.00	78.90	27.40	4.10	7,775.85	0.00	440.00
46	Italy	11.00	0.00	19.05	0.00	8.58	75.25	12.35	26,955,337,473.00	65.60	23.80	1.30	1,975.10	257,757,921,481.60	288.00
47	Jamaica	36.30	81,255,000.00	0.00	0.00	19.69	59.55	14.10	3,843,738.50	43.20	5.80	0.00	0.00	1,089,797,769.00	59.00
48	Japan	8.20	0.00	0.00	0.00	9.53	87.90	3.85	91,529,336,519.00	93.30	30.50	3.55	5,293.75	877,926,395,223.50	3,504.00
49	Jordan	28.80	2,050,890,000.00	28.00	0.00	0.00	84.20	11.85	78,833,688.00	53.40	4.20	0.00	0.00	6,106,026,090.20	228.00
50	Kazakhstan	7.75	88,320,000.00	29.75	164,240.00	0.00	68.85	4.65	3,195,296,949.50	72.90	13.00	0.20	671.85	18,582,516,079.40	78.00

#	Country	Business Start-up Cost & Procedure						Financing SMEs			Tax Rates & Structure		
		Cost of business start-up procedures (% of GNI per capita)	Start-up procedures to register a business (number)	Time required to start a business (days)	Patent applications, nonresidents	Trademark applications, direct nonresident	Charges for the use of intellectual property, payments (current US$)	Firms using banks to finance investment (% of firms)	Lending interest rate (%)	Foreign direct investment, net (current US$)	Total tax rate (% of commercial profits)	Profit tax (% of commercial profits)	Taxes on goods and services (% value added of industry and services)
51	Korea, Rep.	14.50	3.00	4.00	46,219.00	22604.00	9,831,100,000.00	0.00	3.50	22,597,800,000.00	33.20	18.20	6.25
52	Kuwait	2.15	12.00	31.00	0.00	0.00	0.00	0.00	4.30	5,154,974,475.20	13.00	0.00	0.00
53	Kyrgyz Republic	2.20	4.00	10.00	7.00	2941.00	5,954,773.00	18.40	24.20	-707,451,800.00	29.00	6.40	13.40
54	Latvia	2.55	4.00	9.00	4.00	1857.00	34,403,769.10	7.90	0.00	-626,221,845.10	35.90	6.30	12.65
55	Lebanon	33.70	6.00	15.00	0.00	0.00	34,392,576.10	53.10	7.10	-1,722,151,183.70	30.30	6.10	8.60
56	Lithuania	0.40	2.50	3.50	42.00	1925.00	48,734,343.10	28.00	0.00	-877,594,291.20	42.60	5.90	3.55
57	Luxembourg	2.00	6.00	18.50	90.00	0.00	3,109,206,540.90	0.00	0.00	53,363,428,378.40	20.10	4.20	14.45
58	Macedonia, FYR	0.35	1.50	1.50	0.00	2931.00	56,157,840.00	20.90	7.10	-193,321,690.50	12.90	10.90	15.70
59	Malaysia	6.95	3.00	4.00	6,267.00	19171.00	1,256,980,166.30	36.50	4.60	-1,075,213,447.70	40.00	22.70	3.50
60	Mexico	18.30	6.00	6.30	14,889.00	29425.00	872,540,404.00	16.20	3.40	2,033,154,032.40	51.70	0.00	0.00
61	Moldova	4.45	4.50	5.00	72.00	3945.00	17,870,000.00	20.00	14.10	-265,340,000.00	40.20	9.20	23.00
62	Montenegro	1.50	6.00	10.00	0.00	3488.00	3,445,277.00	29.60	8.90	-687,343,510.10	21.60	8.10	0.00
63	Morocco	9.15	4.50	10.50	742.00	5543.00	95,600,829.90	34.80	0.00	-2,507,734,230.30	49.10	25.40	16.15
64	Namibia	12.10	10.00	66.00	0.00	941.00	9,659,181.20	33.50	9.30	-932,960,816.00	21.30	17.20	9.05
65	Netherlands	4.80	4.00	4.00	288.00	0.00	48,088,044,479.60	0.00	0.00	22,579,600,252.30	41.00	20.40	11.35
66	Nigeria	32.80	8.70	30.80	0.00	0.00	254,687,523.00	6.90	16.80	-3,695,026,235.40	33.30	0.00	0.20
67	Norway	0.90	4.00	4.50	457.00	11453.00	568,362,057.60	0.00	0.00	18,951,620,135.40	39.50	23.60	12.65
68	Oman	2.80	5.00	7.00	0.00	2378.00	0.00	0.00	4.80	32,509,752.95	22.90	11.00	0.00
69	Panama	6.35	5.00	6.00	274.00	5482.00	107,100,000.00	1.20	7.50	-4,510,700,000.00	37.20	12.40	0.00
70	Peru	9.50	6.00	26.00	1,204.00	8877.00	292,156,767.90	45.90	16.10	-6,734,433,651.20	35.90	22.70	8.90
71	Philippines	16.35	16.00	31.50	3,255.00	11240.00	612,611,626.00	12.10	5.60	-122,280,021.00	42.90	20.30	4.45
72	Poland	12.55	4.00	30.00	155.00	3275.00	2,432,000,000.00	30.60	0.00	-4,438,000,000.00	40.30	14.50	13.10
73	Portugal	2.25	3.00	2.50	18.00	2201.00	691,011,001.30	0.00	0.00	2,106,462,279.40	41.00	13.60	13.65
74	Puerto Rico	1.30	6.00	6.00	0.00	0.00	0.00	0.00	0.00	0.00	65.80	25.90	0
75	Qatar	5.15	8.00	8.50	477.00	6203.00	0.00	0.00	4.40	2,952,472,527.50	11.30	0.00	0.00
76	Romania	2.05	5.00	8.00	84.00	2326.00	808,229,990.80	28.00	6.80	-3,077,696,841.20	42.00	10.90	14.45
77	Russian Federation	1.15	4.40	10.85	16,236.00	26903.00	5,633,850,000.00	11.30	15.70	15,710,010,000.00	47.00	0.00	7.10
78	Saudi Arabia	4.05	12.00	19.75	135.00	0.00	0.00	0.00	0.00	-2,620,698,333.30	15.00	2.20	0.00
79	Serbia	6.70	6.00	12.00	10.00	5175.00	180,302,819.60	27.00	0.00	-2,000,629,971.60	39.70	16.00	21.15
80	Singapore	0.60	3.00	2.50	9,009.00	15870.00	17,284,745,331.20	0.00	5.40	-29,777,462,586.10	18.40	2.00	4.60
81	Slovak Republic	1.50	6.50	11.50	23.00	2122.00	521,977,872.80	37.40	0.00	-991,834,925.80	51.20	10.50	10.70
82	Slovenia	0.00	2.00	6.00	0.00	1402.00	221,333,711.90	30.40	0.00	-1,370,123,685.20	31.00	12.70	15.35
83	South Africa	0.30	6.00	46.00	6,750.00	14943.00	1,708,385,694.10	0.00	9.40	3,551,758,122.00	28.80	21.70	11.35
84	Spain	5.25	7.00	14.00	225.00	4122.00	4,519,625,973.30	0.00	0.00	26,636,733,691.40	50.00	13.30	8.05
85	Sri Lanka	19.70	8.50	10.50	0.00	0.00	0.00	43.60	8.00	-627,976,968.50	55.20	1.10	6.70
86	Swaziland	23.35	12.00	30.00	0.00	679.00	63,588,841.30	0.00	9.00	-27,594,867.15	34.70	25.60	0.00
87	Sweden	0.50	3.00	11.50	441.00	1983.00	4,203,280,258.00	17.90	0.00	10,906,442,335.50	49.10	13.10	14.25
88	Switzerland	2.00	6.00	10.00	568.00	17332.00	12,329,972,403.90	0.00	2.70	2,005,482,615.20	28.80	9.30	5.60
89	Tajikistan	22.40	4.00	19.50	0.00	2241.00	65,600.00	13.20	25.80	-391,248,980.00	81.80	17.70	0.00
90	Thailand	6.50	6.00	27.50	6,924.00	18144.00	4,125,381,683.00	17.00	6.60	3,497,151,473.60	27.50	19.50	9.60
91	Trinidad and Tobago	0.70	7.00	11.50	184.00	1106.00	0.00	36.70	8.20	-352,550,000.00	32.20	21.90	5.80
92	Tunisia	4.05	10.00	11.00	400.00	3534.00	16,551,380.50	22.90	0.00	-1,031,645,365.65	59.90	12.80	10.80
93	Turkey	16.50	8.00	7.50	331.00	14490.00	682,000,000.00	44.20	0.00	-7,116,500,000.00	40.90	17.90	17.45
94	Ukraine	0.90	5.00	14.00	2,356.00	12102.00	358,000,000.00	30.30	21.80	-3,012,000,000.00	52.20	9.00	16.50
95	United Arab Emirates	6.25	6.00	8.00	1,447.00	13329.00	0.00	0.00	0.00	0.00	15.90	0.00	0.40
96	United Kingdom	0.20	5.00	5.25	7,844.00	8174.00	12,426,547,510.10	0.00	0.00	22,799,643,393.30	32.00	19.20	13.30
97	United States	1.15	6.00	5.60	293,706.00	58672.00	39,495,000,000.00	0.00	3.30	126,864,500,000.00	43.90	0.00	0.60
98	Uruguay	22.55	5.00	6.50	639.00	3597.00	48,483,574.20	13.70	15.80	-1,613,885,085.00	41.80	23.60	15.00
99	Vietnam	5.10	10.00	27.00	3,960.00	12181.00	0.00	30.30	7.10	-7,497,000,000.00	39.40	14.50	0.00
100	Zambia	32.75	6.00	7.50	25.00	2145.00	1,827,525.00	12.20	13.30	-2,087,066,000.00	18.60	2.00	5.05

#	Country	Governmental Regulations & Policies			Labor & Skills					Technology Advancement						Competition
		Exports of goods and services (% of GDP)	Trade in services (% of GDP)	Net official development assistance and official aid received (current US$)	Labor force with tertiary education (% of total)	Secondary education, vocational pupils	Government expenditure on education, total (% of government expenditure)	Wage and salaried workers, total (% of total employed)	Unemployment, total (% of total labor force)	High-technology exports (current US$)	Internet users (per 100 people)	Fixed broadband subscriptions (per 100 people)	Research and development expenditure (% of GDP)	Researchers in R&D (per million people)	Manufacturing, value added (current US$)	Listed domestic companies, total
51	Korea, Rep.	45.90	15.30	0.00	0.00	374,676.00	0.00	72.20	3.30	131,953,914,182.00	89.90	40.20	4.20	6,677.80	369,570,020,090.00	1,948.00
52	Kuwait	69.40	26.50	0.00	19.10	6,201.00	0.00	96.80	3.10	44,704,387.50	82.10	1.40	0.20	127.65	9,947,541,250.10	198.00
53	Kyrgyz Republic	39.85	27.60	580,245,000.00	18.20	49,208.00	18.18	54.95	8.20	16,198,580.00	30.20	3.70	0.10	0.00	793,441,417.50	26.00
54	Latvia	58.80	25.90	0.00	32.55	28,440.00	8.84	88.40	10.95	961,867,110.00	79.20	25.10	0.65	1,842.90	3,373,218,683.05	0.00
55	Lebanon	56.90	62.40	720,315,000.00	0.00	57,689.00	7.84	0.00	6.30	39,722,416.50	74.00	22.80	0.00	0.00	3,958,806,709.30	10.00
56	Lithuania	77.30	27.40	0.00	38.80	28,822.00	0.00	88.05	11.55	1,773,594,152.00	71.40	27.80	1.00	2,924.35	8,293,505,778.15	0.00
57	Luxembourg	213.80	286.40	0.00	45.40	0.00	0.00	91.00	6.00	673,934,496.00	97.30	36.50	1.30	4,585.90	2,478,318,731.30	27.00
58	Macedonia, FYR	48.50	26.40	204,000,000.00	20.55	0.00	0.00	72.60	28.45	110,003,537.00	70.40	17.20	0.35	757.40	1,083,625,310.80	32.00
59	Malaysia	71.00	25.20	13,645,000.00	21.45	0.00	19.69	74.50	2.60	57,257,697,127.00	71.10	9.00	1.20	1,922.60	67,548,585,328.40	892.00
60	Mexico	35.30	4.80	684,080,000.00	23.30	0.00	19.20	66.65	4.90	45,780,911,356.00	57.40	11.60	0.50	317.15	195,791,498,962.00	136.00
61	Moldova	43.40	27.60	432,215,000.00	25.05	32,781.00	21.39	68.05	4.25	17,928,258.00	49.80	15.50	0.40	647.95	775,853,633.60	0.00
62	Montenegro	43.30	45.50	109,940,000.00	24.65	0.00	0.00	83.00	19.30	6,492,297.00	64.60	18.10	0.40	660.05	192,258,936.00	129.00
63	Morocco	33.50	22.10	2,125,535,000.00	9.05	0.00	17.81	43.55	9.70	909,881,681.00	57.10	3.40	0.70	854.40	17,317,031,446.40	74.00
64	Namibia	43.90	16.70	243,635,000.00	6.70	0.00	26.19	63.30	18.80	63,599,468.00	22.30	1.70	0.10	0.00	971,537,307.90	8.00
65	Netherlands	82.80	38.90	0.00	33.15	565,413.00	11.69	83.60	6.80	69,673,950,438.50	93.10	41.70	2.00	4,519.65	83,150,715,315.50	100.00
66	Nigeria	18.20	4.65	2,495,640,000.00	0.00	0.00	0.00	0.00	7.50	111,834,296.50	47.40	0.00	0.00	0.00	45,357,020,685.10	183.00
67	Norway	37.10	22.30	0.00	39.95	130,789.00	15.16	92.90	3.45	4,729,878,109.00	96.80	38.90	1.70	5,624.35	28,587,826,805.80	171.00
68	Oman	72.35	16.30	0.00	0.00	109.00	0.00	96.30	7.25	195,409,499.00	74.20	5.60	0.20	148.80	8,180,754,226.25	116.00
69	Panama	57.05	30.20	28,960,000.00	32.75	0.00	13.02	67.50	4.20	141,177.00	51.20	7.90	0.10	38.65	2,892,497,500.00	30.00
70	Peru	21.00	7.20	342,310,000.00	23.35	44,270.00	14.79	47.55	4.10	195,064,245.00	40.90	6.40	0.00	0.00	28,964,846,767.00	212.00
71	Philippines	27.90	17.80	433,080,000.00	27.20	0.00	0.00	57.40	7.10	26,192,288,870.00	40.70	3.40	0.10	137.25	58,669,403,374.10	262.00
72	Poland	49.40	16.00	0.00	30.30	785,567.00	11.09	78.40	9.80	13,444,666,348.00	68.00	19.50	0.90	1,943.25	81,337,047,353.80	872.00
73	Portugal	40.30	21.10	0.00	21.80	219,620.00	10.34	79.20	15.35	1,902,608,832.00	68.60	29.60	1.30	3,657.50	23,336,760,563.40	47.00
74	Puerto Rico	74.30	0.00	0.00	0.00	7,204.00	0.00	84.55	14.50	0.00	79.50	18.00	0.40	260.10	47,163,395,081.50	0.00
75	Qatar	55.40	27.40	0.00	0.00	570.00	0.00	99.40	0.30	30,612.00	92.90	10.10	0.50	597.10	16,495,604,395.60	43.00
76	Romania	41.10	16.60	0.00	18.10	0.00	8.39	67.55	7.15	3,536,097,090.00	55.80	19.80	0.40	930.00	36,463,560,136.60	82.00
77	Russian Federation	29.50	10.60	0.00	55.90	1,547,945.00	0.00	92.90	5.30	9,249,223,001.50	73.40	18.80	1.15	3,087.35	69,975,616,789.00	251.00
78	Saudi Arabia	33.70	16.20	0.00	0.00	177,014.00	0.00	94.60	5.65	272,788,564.00	69.60	12.00	0.00	0.00	79,457,333,333.30	171.00
79	Serbia	47.70	23.80	575,695,000.00	21.00	212,233.00	10.10	68.60	22.15	354,236,651.00	65.30	16.80	0.75	1,422.85	6,745,563,523.75	775.00
80	Singapore	176.50	96.70	0.00	39.65	0.00	20.43	84.90	2.90	136,485,323,909.00	82.10	26.50	2.10	6,661.85	53,049,292,818.80	483.00
81	Slovak Republic	93.80	18.50	0.00	20.70	155,326.00	9.81	84.55	13.75	6,720,319,370.00	85.00	23.30	0.85	2,718.05	16,305,755,794.60	68.00
82	Slovenia	77.80	26.00	0.00	30.50	61,245.00	12.08	82.25	9.85	1,432,654,148.00	73.10	27.60	2.50	4,183.35	8,501,627,555.40	46.00
83	South Africa	30.90	9.80	1,182,940,000.00	16.65	359,191.00	19.89	85.70	24.85	2,352,168,262.50	51.90	5.30	0.70	394.90	36,284,946,194.40	316.00
84	Spain	33.10	15.20	0.00	36.75	547,099.00	9.95	82.25	25.50	14,240,904,065.00	78.70	28.30	1.25	2,647.50	145,013,862,703.80	3,623.00
85	Sri Lanka	20.50	12.70	444,665,000.00	17.50	149,818.00	9.04	55.10	4.50	59,295,119.00	30.00	3.10	0.10	108.40	14,921,914,161.20	294.00
86	Swaziland	51.45	20.90	100,945,000.00	0.00	0.00	20.32	0.00	22.30	0.00	30.40	0.50	0.00	0.00	1,474,759,689.30	0.00
87	Sweden	45.20	26.60	0.00	36.05	245,912.00	13.30	89.55	8.05	14,933,994,823.00	90.60	36.10	3.25	6,769.05	73,262,556,130.00	0.00
88	Switzerland	63.50	30.40	0.00	37.05	0.00	16.00	84.80	4.45	53,257,659,389.00	88.00	44.80	3.00	4,481.10	119,468,548,125.70	234.00
89	Tajikistan	20.35	9.50	372,865,000.00	0.00	0.00	15.46	0.00	11.05	0.00	19.00	0.10	0.10	0.00	775,037,398.15	0.00
90	Thailand	68.55	28.30	188,680,000.00	14.95	729,843.00	21.35	42.55	0.80	34,446,805,197.00	39.30	9.20	0.40	758.75	114,338,139,877.45	639.00
91	Trinidad and Tobago	34.70	25.75	0.00	0.00	0.00	0.00	77.65	3.80	3,102,673.00	69.20	20.70	0.10	0.00	1,445,730,234.90	0.00
92	Tunisia	46.65	17.35	815,845,000.00	18.80	99,345.00	23.42	71.70	13.30	592,717,059.50	48.50	4.30	0.65	1,797.85	7,485,894,364.15	78.00
93	Turkey	28.00	9.60	3,142,270,000.00	19.40	2,269,651.00	0.00	65.10	8.95	2,323,079,468.00	53.70	12.40	0.95	1,162.55	111,925,661,029.40	392.00
94	Ukraine	52.80	25.50	1,093,580,000.00	45.80	271,345.00	13.57	82.50	7.45	2,055,514,345.00	49.30	11.80	0.75	1,095.80	10,943,890,280.00	158.00
95	United Arab Emirates	99.65	0.00	0.00	0.00	6,443.00	0.00	0.00	3.70	1,006,737,411.00	91.20	12.80	0.60	0.00	35,119,400,953.00	125.00
96	United Kingdom	27.40	19.50	0.00	39.55	2,083,049.00	12.84	84.65	6.90	69,340,644,491.00	92.00	37.70	1.70	4,219.05	260,516,656,960.80	1,858.00
97	United States	12.60	6.90	0.00	33.50	0.00	12.97	93.30	6.80	154,353,963,992.00	74.60	31.50	2.70	4,014.70	2,031,739,000,000.00	4,381.00
98	Uruguay	22.30	10.60	62,345,000.00	21.90	0.00	14.93	72.50	6.80	180,184,158.00	64.60	26.30	0.30	524.30	7,146,857,607.40	0.00
99	Vietnam	89.80	13.80	4,150,535,000.00	0.00	0.00	21.19	34.75	2.25	29,341,628,918.00	52.70	8.10	0.20	0.00	26,512,906,238.20	307.00
100	Zambia	37.00	9.20	1,068,385,000.00	0.00	0.00	0.00	20.60	13.20	29,222,469.50	21.00	0.10	0.00	0.00	1,710,060,332.00	20.00

CURRICULUM VITAE

Suhail H. Serbaya

Place of birth: Makkah, Saudi Arabia

Education

- BSc in Industrial Engineering, King Abdul Aziz University, Jeddah, Saudi Arabia
  Senior Project Title: "Effectiveness of training system in King Abdul Aziz University & King Fahad University".
- MSc in Industrial Engineering, King Abdul Aziz University, Jeddah, Saudi Arabia
  Thesis Title: "Studying the Impact of Implementing the Incubator Concept on Small-to-Medium Enterprises".
- PhD in Industrial Engineering, University of Wisconsin-Milwaukee, WI, USA
  Dissertation Title: "Facility Location Decision For Global Entrepreneurial Small-to-Medium Enterprises Using Similarity Coefficient-Based Clustering Algorithms".

Affiliations/Memberships

- Saudi Society for Systems and Industrial Engineering (SSSIE)
- Badir Program for Technology Incubators (badir) – Founding member
- Business & Knowledge Alliance executive board
- Business & Knowledge Alliance advisory board
- Arab Trade Center group

Teaching Experience

- Instructor and co-designer of business incubation program at Business & Knowledge Alliance (King Abdul Aziz University)

- Coaching university students towards identification of valid business opportunities
- Mentoring university students towards preparation of business plans and development of associated feasibility studies for their small ventures
- Designing, developing and conducting training programs on New Venture Creation
- Assisting in developing the curriculum courses of Institute of Business Administration (IBA) (presently College of Business Administration (CBA))

Research Experience

- Coordinating scientific & financial affairs of Consultation Services provided by university's faculty for both governmental & private sectors.
- Appraising, evaluating and monitoring research funding and collaboration

Awards/Honors

- Chancellor's Graduate Student Award (Academic Year 2012-2013), University of Wisconsin-Milwaukee, WI, USA

Presentations

- Work paper "Technology absorption and developing a business plan for innovators and inventors", June 2004, Conference of Innovators and Inventors, King AbdulAziz University – Jeddah, Saudi Arabia
- Work paper "Reinforcing the quality of tourism industry via SMEs", December 2004, Conference of Innovative Tourism, Saudi Arabian Quality Council – Abha, Saudi Arabia
- Work paper "The need for developing a strategy for SMEs promotion in Saudi Arabia", April 2005, Conference of SMEs Development in Arabian Countries, the Islamic Development Bank – Jeddah, Saudi Arabia