September 2017

Introduction of Similarity Coefficient-based Clustering Algorithms to Global Petrochemical Facility Location

Ali Saeed Alarjani
University of Wisconsin-Milwaukee

INTRODUCTION OF SIMILARITY COEFFICIENT-BASED CLUSTERING ALGORITHMS TO GLOBAL PETROCHEMICAL FACILITY LOCATION

by

Ali AlArjani

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Engineering

at

The University of Wisconsin-Milwaukee

December 2017

ABSTRACT

INTRODUCTION OF SIMILARITY COEFFICIENT-BASED CLUSTERING ALGORITHMS TO GLOBAL PETROCHEMICAL FACILITY LOCATION

by

Ali Saeed AlArjani

The University of Wisconsin-Milwaukee, 2017
Under the Supervision of Associate Professor Nidal Abu Zahra

This research introduces a similarity coefficient-based clustering algorithm to determine the best location for a petrochemical manufacturing facility. The most global petrochemical critical attributes have been selected from relevant literature about manufacturing activities. These critical attributes have been quantified by real world numbers from the World Bank database and have been employed in the proposed model of the research. The model of the research uses the selected critical attributes data and clusters a hundred countries in similar groups according to their attractiveness level to the petrochemical facility location.

The outcomes of the developed model are classifications that show the potential country for locating a petrochemical facility. Moreover, all countries have been ranked first according to their high potential cluster and within each cluster. These rankings also help to distinguish the candidate countries assigned to the same cluster.

The flexibility and the capacity of the developed model give higher advantages over the other facility location solution models. This research suggested a new quantitative petrochemical facility location selection criteria model that cluster locations in groups based on their similarity and dissimilarities by analyzing the data of the selected attributes.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First, I must express my very profound gratitude to my parents, my brothers, my sisters and to my wife for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.

I would like to express my sincere gratitude to my advisor Prof. Nidal Abu Zahra for the continuous support of my Ph.D study, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. The door to Prof. Nidal Abu Zahra office was always open whenever I had a question about my research or writing. He consistently steered me in the right the direction whenever he thought I needed it. I could not have imagined having a better advisor and mentor for my Ph.D study.

I would also like to acknowledge Prof. Hamid Seifoddini and I hope he feels better and got health back normal soon, I am gratefully indebted to his for his very valuable comments on this thesis.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Prof. Wilkistar Otieno, Prof. Matthew Petering and Dr. Xiaohang Yue, for their insightful comments and encouragement.

Author

Ali AlArjani

CHAPTER ONE

Introduction

1.1 Problem Background

The manufacturing industry around the globe relies on certain critical attributes based on the nature of the business when considering new locations. Locating these industries in any country will develop growth and improve its economy. The methods and theories used to select the best location are highly important for the industry and will provide a tool that can be used for both short- and long-term strategies, including future expansion.

The type of the industry determines the critical attributes of a decision of future locations because each industry has its own priorities in terms of the objectives, for example, labor costs, sales, tax system, etc. The algorithm introduced in this thesis provides decision makers with the opportunity to consider a large number of attributes for deciding the most appropriate location for a new petrochemical facility. This model will give a set of options that offer multiple locations for the decision makers to select.

The globalization of the petrochemical industry raises the need of relocating and establishing new plants around the globe. Recent studies show how the international open market growth trend caught the attention of researchers to propose appropriate approaches other than the classical decision-making methods that are based on old defined critical attributes such as natural resources. Embracing the new critical factors on the global market will change the process for making decision about a new petrochemical facility location.

These new critical factors such as the high technology and environmental regulations along with geopolitical factors are very critical attributes for the petrochemical industry. The new factors influence the sustainability of the petrochemical site in general. Those factors also differ from one country to another, which affects the petrochemical facility location.

Due to the change in the demand and many other reasons, the petrochemical industry is expanding and shifting from the west to Asia and the Middle East. The fastest growing areas in petrochemicals are the Middle East and Asia outside of Japan with double digit growth rates. These areas have started investing tremendous amount of money on building their own facilities that are capable of converting the crude oil to basic chemicals that make up the petrochemical industry.

In the United States, as well as in most developed countries, investments will be spent on the existing facilities to meet environmental regulations rather than building green sites. The current analysis shows that billions of dollars in the petrochemical industry were spent on the existing facilities around the globe and most of these investments were used to increase the productivity of the existing plant. With the increasing of the globalization in this industry fewer new petrochemical plants will be built in United States and most developed nations as well. Also, another result of globalization is pushing towards partnerships and joint venture that also lead to fewer players in the global market place.

In these rapid global changes, many petrochemical plants shift or expand due to many different reasons or factors. Corporate leadership needs to conduct intensive studies before deciding where to expand or locate a petrochemical facility. The decisions involve a combination of factors. Some of these are under their control and can be manipulated but for other factors, regulations and rules need to be followed. These changes raise the need for more effective research methods to facilitate the global location decisions.

Hierarchical clustering analysis was used in this research to analyze a set of factors versus a set of countries that could serve as locations for new petrochemical plants. The countries' classifications, based on similarities and dissimilarities, could be executed by several existing methods. In data mining clustering analysis has been used in the previous research more often than the other methods due the reasonable outcomes it yields. The algorithm developed in this research has higher flexibility to add or remove factors at any stage of the analysis without reworking the whole model setup. This feature is not available in any other available facility location solution models.

The flexibility and the capacity of the developed model give higher advantages over the other facility location solution models. This research suggested a new quantitative petrochemical facility location selection criteria model that cluster locations in groups based on their similarity and dissimilarities by analyzing the data of the selected attributes (factors).

1.2 Problem Statement

The use of clustering analysis algorithms has not been introduced to the petrochemical facility location problem until this research model was developed. There is a need of a facility location approach that considers the flexibility, capacity and quantifying the attributes. This approach gives an opportunity of using real numbers from the World Bank database website in determining the best location for a new petrochemical facility.

Quantifying the critical factors in this model has an absolute advantage over the previous methods of solving facility location problems. Moreover, previous approaches for solving facility location problems had limited flexibility of adding or removing some factors in the middle of the study, which complicate the decision in many cases.

1.3 Significance of Research

The developed model will open a new horizon for decision makers around the globe by analyzing multiple critical factors for a new petrochemical facility location. The significance of this research can be as follows:

- Defining the critical factors that were the most frequent addressed attributes in the literature review for petrochemical industry locations and considering recent changes in the environment and transportation regulations.

- Introducing similarity coefficient clustering for determining appropriate petrochemical facility locations. This hybrid model consists of three main functions: clustering, quantifying critical factors and ranking the potential location based on the selected factors.

- Ranking the potential location (countries) within their clusters based on defined attributes and weights assigned to each attribute.

- This proposed model provides the first petrochemical facility location selection criteria that quantify the critical factors that influence the petrochemical industry locations by using real numbers from the World Bank website.

- This research model gives the petrochemical industry leaders and decision makers an opportunity to use multiple attributes (factors) versus multiple objects (locations) while the previous research dealt with either multiple attributes versus single, double and triple objects or the other way around.

1.4 Research Objectives

The objectives of this research are as follows:

1. To define the critical attributes to be considered in the assessment and selection of a global petrochemical facility location according to the desired objectives.
2. To develop a quantitative model based on similarity a coefficient-based clustering algorithm that quantifies the critical attributes by real world number for the selection of a global petrochemical facility location.

   To rank the locations based on their attractiveness level to petrochemical facility location within each cluster and in general to check which cluster is a higher potential. This research provides a new flexible model that gives a variety of options to the petrochemical industry leaders to make their decisions based on quantified factors by real numbers and includes and excludes some factors according to their individual needs. .

CHAPTER TWO

2.1 Similarity coefficient-based clustering background

(Mcauley 1972) implemented the similarity coefficient of (Jaccard 1908), whose contribution was in manufacturing systems, by defining the similarity coefficient on how similar machines by using the number of parts visiting both machines and the number of parts visiting either machine.

In another contribution of (Sokal and Sneath 1961) a similarity coefficient was explained in a more comprehensive way and the same authors in (Sneath and Sokal 1973) developed the similarity coefficient qualification of the similitude between the parameters in two groups of data matrices that represent the indications of the states of the two taxonomic systems.

(Anderberg 1973) had a slightly different approach by involving the manufacturing data such as processing order and production sizes, which had not been examined before. Then a huge step was accomplished by Seifoddini in (H. K. Seifoddini 1989) to fill the gap of using a similarity coefficient that engaged the production volume and sensitivity between the machines, which made the coefficient more useful in actual practice.

Seifoddini did not touch on the relationship between the process sequence and a similarity coefficient but he did acknowledge that it was an important relationship. Seifoddini went a step further with a corporation called Djassemi when they adopted a Jaccard similarity coefficient (JSC) to be more flexible and overcome the issues of production data (H. Seifoddini and Djassemi 1991; Seifoddini, Hamid 1995).

A new performance measure was conducted by a grouping capability index (GCI) (Seifoddini & Hsu, 1994). That measure has been heavily used in subsequent research. In that study three similarity coefficients were tested.

Seifoddini (1988) as well as Gupta and Seifoddini (1990) proposed the advantages of using a similarity coefficient in these points:

- Simplicity of application
- Flexibility of analyzing the manufacturing data of a cell formation process
- Ease of implementation in computer software.
- Ability to adopt constraints because the similarity coefficient method gives a set of alternative solutions.

2.2 Global facility location background

In global manufacturing, deciding the location of the manufacturing plant is a strategic key factor that could shape the success of the manufacturing firm among the competitors around the globe (Maccarthy and Atthirawong 2003). (Tomback 1995) called the global location context a game of timing. Embracing a decision of building a new business location or expanding the existing facility has long term commitments financially and for allocation of human resources (Epping 1982). An increase in The number of firms planning on making a global basis location (Flaherty 1996) has increased over the last two decades.

In the past, it was very normal for any manufacturing industry to stay in the same location for decades. But the best location for a certain sector of business today might not the best location next year (Epping 1982) because of the rapid change in cost and demand (Lösch

1939). This means that if a firm does not respond to these rapid changes it will go out of business sooner than competitors no matter how good it was.

Over a hundred plants owned by US firms located in Asia, Latin America and Europe were surveyed and showed these driving factors addressed in the management decision process of investing abroad. The factors include the nature of the hosting government, the environment, accessibility, the area reputation, industrialization, basic service availability, policies, site cost, host tax and incentives, and labor and staff availability (Bass, Mcgregor, and Walters 1977). There are many factors that can be involved in deciding a new global manufacturing location (McCarthy, 2003)

Thunen began an early economic analysis of an industrial location process theory that was based on the approach of least-cost (Thunen 1875). Then (Launhardt 1885) had a significant contribution in considering industrial location analysis from a demand and cost prospective. He also pointed out the transportation cost as another critical factor.

Weber in 1909 devised a more comprehensive theory for manufacturing plant location (Weber 1929; Isard 1956). Three critical factors were addressed: labor and transportation cost and, as Weber called it, the agglomeration force of the firms. Weber’s theory was used in many later studies for the sake of better understanding of the decision process (Tellier and Vertefeuille 1995).

Hotelling’s contribution in 1929 was significantly important in the historical part of the industrial location, which became a base of much later research in the industrial location analysis process. Hotelling had produced two main points: first, he linked competition to the location decision process and second, he developed a tendency of the companies to make their location close to their market, which Weber in 1909 had called agglomeration force factors (Hotelling 1929).

(Lerner and Singer 1937) extended Hotelling's work but they stated that Hotelling's theory was not always applicable to the tendency of firms making a location decision. Three location factors were determined: how much the buyer was willing to pay, the size of the market, and the transportation cost (Lerner and Singer 1937). An interdependence model of one dimensional bounded was examined by (Balvers and Szerb 1996; Katz 1995; Smithies 1941; Bertil Ohlin 1935; Bertil Ohlin 1952).

In 1939 a maximum-profit theory was developed by August Lösch. Lösch's contribution was a location analysis that considered a free economy to select the firm's site according to the cost and demand curves (Lösch 1939). Lösch's approach for industrial location was integrated into a cost and demand theory by (Hoover 1937; Hoover 1948). Hoover emphasized that the link between the transportation cost and the firm location is not proportional. A new plant location theory based on cost and demand attributes was developed by Greenhut in 1956. An impressive contribution was made by Button in 1996, who continued Greenhut's ideas about the industrial location theory and its own economic factors (Button 1996).

Throughout the literature review, the surveys and questionnaires methods are greatly valued for the accuracy of the data they provide and the sense of reality they offer. Surveys have been used in the manufacturing field, as well as other research applications, for many purposes in the decision-making process, data mining, and so on. Surveys have their unique advantages for many valid reasons; such as the ability of shaping the survey to targeted purposes where the input data serve the study goals directly.

2.2 International facility location factors

A number of studies conducted on the critical factors concerning international location decisions focused on a limited scope of manufacturing operations (Maccarthy and Atthirawong 2003; Siebert 2006). The literature on the international industrial location is divided into two categories, either developing theoretical concepts or empirical studies. A common feature of those two types is strongly recommending that global investors realize any type of the host government reactions, which tend to be very sophisticated (Vernon 1968; Vernon 1971). To some extent Tomback called the international location decision a game of timing (Tomback 1995).

Horst conducted a survey of 1191 manufacturing firms with many locations around the globe. He studied the impact of the investment process of firms that made their direct investment in Canada and the ones that did not and also the firms with branches around the globe and the ones with only one domestic location (Horst 1972).

Another researcher who investigated firms and industry critical factors was Vernon in 1971. He studied 187 U.S. manufacturing firms with at least six subsidiaries or more and came up with a set of influencing factors for these firms. Another study of the process engaged by multinationals to address the political risk was conducted by Rummel and Heenan in 1978. (Rummel and Heenan 1978) described the perspective of the host critical factors that affect the decision process of the international industrial location. This study stated factors that included the economic climate, domestic instability, the political climate and foreign conflict (Rummel and Heenan 1978).

Incomplete and inadequate research information on the decision process of the industrial location could lead to huge and costly failures. In addition, marginal non-economic factors could also have a tremendous impact on the location decision that had been addressed adequately in the literature (Piper 1971). Similarly, political and social factors are significantly important, which has led some multinational enterprises to minimize their risk by locating their manufacturing plants in a different category of countries (Vernon 1968; Vernon 1971; Belli 1971).

Many authors considered political instability a high risk factor due the complexity of creating a clear view of the hosting government policies (Annett 2001; Smith-Hamilton, A. & Omar 2005). In another study personal views could affect the global location decision (Bass, Mcgregor, and Walters 1977)

A number of researchers consider that the firm location decision is a high strategic level for the long run (Vastag, Gyula, Sándor Kerekes 1996). Other research determined that the firm has to be located near enough to its competitor, which places it in the same market location (Venables 1996). What are called soft factors in the location decision process has significant weight, according a Dziembowska-Kowalska and Funk study in 2000. The “cultural multiplier” was used as a measure to check the flow of income resulting from such activities (Dziembowska-Kowalska, J. & Funck 2000).

Over a hundred Spanish firms with subsidiaries around the globe contacted in a study for new locations in Latin America responded with new factors that were not included for different locations (Galan, J. & Gonzalez-Benito 2006).

The literature reviews of the critical factors for the global industrial location were gathered from many studies. The primary factors revealed in the literature included:

- Transportation-related factors
    - Availability of airway facilities
    - Availability of highway facilities
    - Availability of railroad facilities
    - Availability of trucking services
    - Availability of water (port) transportation
    - Availability of pipeline facilities
    - Cost of raw material transportation
    - Cost of finished goods transportation
    - Availability of postal services

(Hoover 1937; Chisholm 1971; Lowe, J. & Moryadas 1975; Losch 1954; Moriarty 1980; Greenhut 1956; McKinnon 1983; Mckinnon 1989; Gold 1991; Thisse, J. & Wildasin 1995; McMillan 1965; Bater, J. & Walker 1977; Pietlock 1992)

- Labor-related factors
    - Availability of skilled labor
    - Wage rates
    - Availability of unskilled labor
    - Existence (or non-existence) of labor unions
    - Educational level of labor
    - Dependability of labor

- Availability of male labor
- Availability of female labor
- Cost of living (housing)
- Worker stability

(Greenhut 1956; Rees 1972; Friedman 1977; McMillan 1965; Sant 1975; Moriarty 1980; Schmenner 1982; Dicken, P. & Lloyd 1978; Malecki 1984; Saxenian 1985; Haitani, K. & Marquis 1990; P. Dicken 1986; Noyelle, T. & Stanback 1984; Wheeler, D. & Mody 1992; Townroe 1969; Carnoy 1972; Pietlock 1992; Coughlin, C.; Joseph, V. & Vachira 1990)

- Raw materials-related factors
  - Availability of raw materials (or components)
  - Proximity to materials and components
  - Availability of storage facilities
  - Location of suppliers
  - Cost of freight (of raw materials and components)

(Weber 1929; McMillan 1965; Moriarty 1980; Schmenner 1982; Wheeler, D. & Mody 1992; Greenhut 1956)

- Market-related factors
  - Proximity to consumers' goods markets
  - Proximity to producers' goods markets
  - Anticipation of growth of markets
  - Shipping costs to market areas
  - Availability of marketing services

- Attainment of favorable competitive position
- Income trends
- Population trends
- Consumer characteristics
- Location of competitors
- Future expansion opportunities
- Size of market
- Industrial site

(Hotelling 1929; Hoover 1948; Greenhut 1956; Moriarty 1980; Saxenian 1985; Mckinnon 1989; Chisholm 1971; Schmenner 1982; Wheeler, D. & Mody 1992; Losch 1954; Carnoy 1972; Walters, B. & Wheeler 1984; Pietlock 1992)

- Industrial site-related factors
    - Cost of industrial land
    - Cost of developed industrial park (or area)
    - Acreage (or space) required
    - Availability of space for future expansion
    - Insurance rates (cost of insurance)
    - Availability of lending institutions (such as banks)
    - Proximity to other industries

(Hoover 1948; McMillan 1965; Chisholm 1971; Moriarty 1980; Schmenner 1982; Greenhut 1956; Wheeler, D. & Mody 1992; Bater, J. & Walker 1977; Coughlin, C.; Joseph, V. & Vachira 1990)

- Utilities-related factors
    - Adequacy of water supply

    - Quality of water
    - Cost of water
    - Availability of disposable facilities for industrial waste
    - Availability of fuels
    - Cost of fuels
    - Availability of electric power
    - Cost of electric power

(Greenhut 1956; Moriarty 1980; Schmenner 1982; Gold 1991; McMillan 1965; Bater, J. & Walker 1977; Walters, B. & Wheeler 1984; Pietlock 1992)

- Government attitude-related factors
    - Zoning codes
    - Compensation laws
    - Insurance laws
    - Safety inspection laws
    - Nuisance and environment pollution laws

(Greenhut 1956; Schmenner 1982; McMillan 1965; Coughlin, C.; Joseph, V. & Vachira 1990; Young 1994)

- Tax structure-related factors
    - Tax assessment basis
    - Industrial property tax rates
    - State corporate tax rates
    - Availability of tax free operations

    - State sales tax

(Greenhut 1956; McMillan 1965; Moriarty 1980; Schmenner 1982; Haitani, K. & Marquis 1990; Coughlin, C.; Joseph, V. & Vachira 1990; Wheeler, D. & Mody 1992; Young 1994)

- Climate-related factors
    - Living conditions
    - Relative humidity
    - Monthly average temperature
    - Air pollution

(Greenhut 1956; McMillan 1965; Moriarty 1980; Schmenner 1982; Haitani, K. & Marquis 1990)

- Community-related factors
    - Availability of universities or colleges
    - Availability of schools
    - Availability of religious facilities
    - Availability of library (information) facilities
    - Availability of recreational facilities
    - Attitude of community leaders towards business
    - Availability of medical facilities
    - Availability of malls (shopping centers)
    - Availability of hotels (motels)
    - Availability of banks and financial institutions
    - Community position on future expansion

(Greenhut 1956; McMillan 1965; Bater, J. & Walker 1977; Moriarty 1980; Rees 1972; Malecki 1984; P. Dicken 1986; Haitani, K. & Marquis 1990; Ballance 1987)

- Political situation of foreign country-related factors
    - Stability of regime
    - Protection of expropriation
    - Type of treaties and pacts
    - Type of military alliances (or with which countries)
    - Attitude towards foreign capital

(Carnoy 1972; Ballance 1987; Dicken, P. & Lloyd 1978; Wheeler, D. & Mody 1992; Young 1994)

- Global competition and survival-related factors
    - Availability of material
    - Availability of labor
    - Market opportunities
    - Availability of foreign capital
    - Proximity to other international markets

(Friedman 1977; Haitani, K. & Marquis 1990; Ballance 1987; Wheeler, D. & Mody 1992; Pietlock 1992)

- Government regulations-related factors
    - Clarity of corporate investment laws
    - Regulations concerning joint ventures and mergers
    - Regulations on transfer of earning out of country
    - Taxation of foreign-owned companies

- Foreign ownership laws
- Allowable percentage of employees who may be foreign
- Prevalence of bureaucratic red tape
- Imposing price controls by government
- Requirements for setting local corporations

(Haitani, K. & Marquis 1990; Wheeler, D. & Mody 1992; Coughlin, C.; Joseph, V. & Vachira 1990)

- Economic-related factors
    - Standard of living
    - Size of per capita income
    - Strength of currency against U.S. dollar
    - Balance of payment status
    - Availability and size of government aids

(Thunen 1875; Dicken, P. & Lloyd 1978; Haitani, K. & Marquis 1990; Ballance 1987; Friedman 1977; Wheeler, D. & Mody 1992)

Another study of the process engaged by multinationals to address the political risk by Rummel and Heenan in 1978 (Rummel and Heenan 1978)provided the perspective of the host critical factors that affect the decision process of the international industrial location. The study stated factors that included the economic climate, domestic instability, the political climate and foreign conflict (Rummel and Heenan 1978).

In Tong's (1979) study 242 manufacturing foreign-owned firms were surveyed for the most critical factors affecting these firms' location decision and found the following:

- Availability of a site
- Labor attitudes
- Nearness to markets
- Transportation services
- Space for expansions

Tong's study also provided a second level of important factors:

- Cost and availability of capital
- Proximity to export markets
- Proximity to operations in third countries
- Proximity to the home country

Twenty-one Japanese and Germany origins firms located in the US were surveyed in a study by Chernotsky in 1983. He proposed two groups of factors: first was that market access, the option of a desirable location and an attractive environment to incoming personnel were the most considered affecting factors. A second group of factors included less attention given by these firms to raw material and finished goods, labor and financial motives (Chernotsky 1983).

In another research study 20 foreign subsidiaries located in the US pointed out an important influencing role to the location decision by the local agencies and the state. Their decision process divided into three main steps:

- To select within a specific geographic region in the United States.
- To have options for two or three states within that region.

- To select the optimum locations among three or four locations proposed in the same state.

An extensive literature review of locating a manufacturing facility in the United States was done by (Jungthirapanich, Chamnong 1995). Their study proposed eight location attributes, listing “market” as the most important and “community environment” as the least important. These eight factors are summarized below:

- Market
  - Proximity to markets
  - Local purchasing power
- Transportation
  - Air transportation
  - Land transportation
  - Water transportation
- Labor
  - Labor force
  - Skilled laborers
  - Work stoppages
- Location cost
  - Land cost
  - Plant construction cost
- Raw material and services
  - Raw material availability
  - Accessibility of business services

- Utilities
    - Fuel Availability
    - Water availability
    - Energy cost
    - Energy capability
- Government policies
    - Taxes
    - Local government aid
    - Support for employment training
    - Government debt
- Community environment
    - Cost of living
    - Education
    - Security
    - Housing availability
    - Health systems
    - Human services
    - Environmental concerns
    - Business climate

In the last decades two aspects concern the manufacturers the most, the cost and globalization. Recently researchers found a link between the performance and the location of the manufacturing plant. These two aspects developed a competitive environment between the manufacturers around the globe (Beckman, Sara Lynn 2008; Rezazadeh and Farahani 2010).

India and China are merging to attract more segments of manufacturing industries (Hanson 2012). On the other hand, in the most developed countries the manufacturing sector is getting smaller and smaller (Peter Dicken 2011). The cost factor has driven many western companies to relocate their manufacturing activities in a less expensive host country (Kinkel 2012).

The literature review revealed no approach similar to our approach in terms of the capability, flexibility and functionality of the developed model. Therefore, the closest approach that could be compared with this model is the analytical hierarchal process (AHP) despite the major differences in the nature of the proposed model in this thesis and AHP.

William Ho described the AHP process as the following: "*The AHP consists of three main operations, including hierarchy construction, priority analysis, and consistency verification. First of all, the decision makers need to break down complex multiple criteria decision problems into its component parts of which all possible attributes are arranged into multiple hierarchical levels. After that, the decision makers compare each cluster in the same level in a pairwise fashion based on their own experience and knowledge. For instance, every two criteria in the second level are compared at each time with respect to the goal, whereas every two attributes of the same criteria in the third level are compared at a time with respect to the corresponding criterion. Since the comparisons are carried out through personal or subjective judgments, some degree of inconsistency may occur. To guarantee the judgments are consistent, the final operation called consistency verification, which is regarded as one of the most advantages of the AHP, is incorporated in order to measure the degree of consistency among the pairwise comparisons by computing the consistency ratio. If it is found that the consistency ratio exceeds the limit, the decision makers should review and revise the pairwise comparisons. Once all pairwise comparisons are carried out at every level, and are proved to be* *consistent, the judgments can then be synthesized to find out the priority ranking of each criterion and its attributes*" (Ho 2008).

The flexibility of AHP is very low compared with our developed model. If one, or a group of attributes, is needed to be removed from consideration of the study in the middle of the process, the whole process needs to be reworked unlike this research model that has tremendous capability of continuing analysis without going back to the point of the process.

Moreover, the AHP is built over the probabilities of the group judgments that add up to form the final decision model while the research model uses real world data from the World Bank database. In AHP, it is complicated to figure out the noise of the decision criteria due to the sophisticated steps of the approach. The AHP also converts the evaluations to numerical values so they can be compared with the numerical values of the problem intended to be solved.

CHAPTER THREE

3.1 The factors influencing facility location decision

In this chapter, the locating or relocating of a petrochemical facility location locally or globally is defined as a comprehensive task. Choosing a location for a petrochemical plant is more sensitive than for other manufacturing sites because of the involved aspects considered in such a step due to the new regulations on the transportation equipped trucking services and the environmental restrictions in the most developed nations. Also, the location decision has an impact on the strategic plans of the company for the long run.

Despite the alternative locations, each location has its own strength in certain factors and weaknesses in others that show the need for a comprehensive tool to study all the critical factors of each location by itself or combined with the other potential locations. Clustering algorithms are considered one of these tools that group the potential locations based on the critical factors chosen and measure the similarities and dissimilarities.

There are several types of clustering algorithms but the common feature among them is that they compare the multiple alternatives based on a selected type of critical factors that are chosen by the researcher. In addition, these clustering algorithms give better results proportional with how well the provided data covered all aspects of the study.

The critical factors are the backbone of this study. If one of the major factors is missed or misrepresented, it will have a major influence on the final results. The facility location problem could be solved by multiple ways according to the stated target so one solution method cannot be applied for all. The critical factors described in the literature review section cover all the local and global manufacturing influencing factors in the location decision. So, the most critical and relevant factors have been chosen for the petrochemical facility location study.

The set of critical factors implemented in this proposed model will show the attractiveness level of the locations towards the petrochemical facility location. There are two categories of the critical factors of this study. The first category includes the critical factors that have been chosen from the literature. The second category includes the corresponding factors in the World Bank indices to the selected critical factors of the literature.

The main factors will be listed first, followed by the sub factors underneath it. The listed critical factors are in sequence with the most important first, as follows:

3.1.1 Economic and market factors

3.1.1.1 Purchasing power

- The market purchasing power effect of the petrochemical facility according to the determined goals.

3.1.1.2 Proximity to the markets

- How far the petrochemical facility from the petrochemical industry markets is, making a difference for the reasons due to the shipping to or from.

3.1.1.3 Stability currency versus US dollar

- Stability of the host country currency versus the United States dollar is considered a valid measure.

3.1.1.4 Potential purchasing market

- The location of a future potential market decides the sustainability of the petrochemical facility location in any country.

3.1.1.5 Marketing services

- For the location of a global manufacturing plant, secondary services are as important as the finished products.

3.1.1.6 Characteristics of the consumers

- Each market in any part of the world has its own characteristics that are based on the raw material used and the desired finished product in that market.

3.1.1.7 Market size

- This factor plays a major role in the location decision because the strategies of the petrochemical facility will be based on it.

3.1.1.8 GDP of the country

- The GDP stands for the gross domestic product and it is an essential measure of the health of any country's economy.

3.1.1.9 GDP per capita

- This factor is another measure of comparing the nations with each other, which represents the GDP divided by the number of people in the country.

3.1.1.10 Host government aids

- These aids also encourage and discourage the petrochemical manufacturing location decision.

3.1.1.11 Competitor location

- The number of competitors of the same business is considered an influencing factor in the petrochemical location decision.

3.1.1.12 Expansion opportunity in future

- This factor needs to be considered for the sustainability and the long run success.

3.1.1.13 Related industries location

- How far are the related industries to the location of the petrochemical plant? It could be critical at some point, which leads the location decision makers to accept or reject some countries based on this factor.

3.1.1.14 Shipping cost from and to the market areas.

- This factor covers any indices that measure the cost of the shipping process beginning with the shipping in the raw material to the shipping out of the finished product to the market buyers.

3.1.2 Labor Factors

3.1.2.1 Labor cost

- The cost of the labor is a major critical attribute for the manufacturing facility location decision.

3.1.2.2 Skilled labor

- Trained labor for certain operations is needed according to how sophisticated is the targeted job

3.1.2.3 Wage rates

- Workers' wages measure how much the worker could get paid per unit of time versus the US dollar.

3.1.2.4 Workers unions

- The union of the laborers in the host country availability is considered a measure in the decision to locate the petrochemical facility location.

3.1.2.5 Educated labor

- The level of the education of the worker is another factor that affects the decision makers in any country to locate the petrochemical facility.

3.1.2.6 Unskilled labor

- This indictor describes a cheaper labor force than the skilled laborer where some operations need to train this type of laborer for a specific procedure.

3.1.3 Transportation Factors

3.1.3.1 Pipeline availability

- This indicator is a measure of the availability of the host country to provide a pipeline transportation system.

3.1.3.2 Air transportation

- Another transportation form is the air transportation system, including the number and location of international airports.

3.1.3.3 Highway transportation

- A measure of the road system network that connects the points of interest of the business chain of the host country.

3.1.3.4 Railroad availability.

- Rail system availability measures the transportation from and to the seaports of the host country.

3.1.3.5 Specific equipped trucking services.

- Petrochemical raw material and finished products require an equipped truck specification according to the new rules in most developed countries.

3.1.3.6 Seaport facilities

- Availability and the number of the seaports are considered influencing factors in the location decision.

3.1.3.7 Availability of postal services

- Postal services availability and capability measure of the host country.

3.1.3.8 Warehousing facilities.

- The availability of chemical warehousing facilities either for the raw materials or the finished products.

3.1.4 Geopolitical Factors

3.1.4.1 Regime stability

- This factor is a major indicator for the strategic planners and the petrochemical facility location studies.

3.1.4.2 Military alliances

- The host country's military alliances.

3.1.4.3 Regime relations with the west

- The level of bond with countries of Western Europe and the United States.

3.1.4.4 Impression in United Nation

- United Nations' standpoint toward the host country

3.1.4.5 Foreign capital encouragement

- The host country's motivated system for foreign capital investment.

3.1.4.6 History of the country

- The host country's history of being politically stable or unstable.

3.1.5 Environmental Factors

3.1.5.1 Air pollution

- Air pollution and gases from the production operations to residential areas.

3.1.5.2 Average temperature

- The average temperature in the host country or region where the petrochemical facility will be located.

3.1.5.3 Environmental rules and regulations of the host country

- This factor has caused many strategic plans of the petrochemical industry to change and relocate in the Middle East and Asia instead of the United States and Europe.

3.1.6 Location Factors

3.1.6.1 Facility construction cost

- This factor is a measure of the cost of constructing a petrochemical facility in the host country.

3.1.6.2 Industrial land cost

- Land cost is a considered factor in the location decision.

3.1.6.3 Insurance cost

- The system of the insurance of the host country is another considered factor in the study of the location.

3.1.6.4 Nearness to the other industries

- How far the petrochemical facility is from related industry for reliability reasons due the shipping to or from.

3.1.6.5 Lending services

- This factor measures how the financial institution facilitates the lending services.

3.1.7 Raw material

3.1.7.1 Available raw material

- The availability of the raw material in the host country.

3.1.7.2 Supplier's location.

- The supplier's location could affect the location decision based on how critical the supplier is to the process.

3.1.7.3 Supplier capability

- This factor measures the capability of the supplier.

3.1.7.4 Raw material shipping cost

- When the raw material is in a different location the shipping back and forth is a considered factor.

3.1.8 Government regulations

3.1.8.1 Taxes

- The tax system of the host country.

3.1.8.2 Rules in wiring the money out of the country

- The restriction on wiring money in and out of the host country is a major factor for the corporate companies to make a location decision.

3.1.8.3 Prices control regulations

- This factor is the pricing system of the host country.

3.1.8.4 Foreign-owned firm tax policies

- A measure of the tax policies towards the foreign-owned company.

3.1.8.5 Percentage of employees who should be citizens of the host country

- The host country’s policies on the percentage of the citizens who should be employed in registered foreign companies.

3.1.8.6 Corporate investment motivated regulations

- This measure is how the rules and regulations of the host country are encouraging to investors.

3.1.9 Utility Factors

3.1.9.1 Water availability

- The water factor is a major requirement for a petrochemical facility in any host country.

3.1.9.2 Quality and cost of the water

- The class of the water matters according to the use of it in the petrochemical operations.

3.1.9.3 Electric power cost

- The electricity cost dollar per electricity unit.

3.1.9.4 Gas availability

- Gas system capability is a major factor of the study.

3.1.9.5 Electric power capability

- The electricity system's capability to reliably connect to the manufacturing petrochemical facility is a major factor of the study.

3.1.9.6 Availability of toxic disposal facilities

- There are toxic gases and disposals from the petrochemical plants that burn in the air but still some gases stay partially toxic in the air.

3.1.9.7 Availability of nuclear and coal facility

- Availability of the nuclear technology affects the cost of the energy used in the petrochemical facility.

3.1.10 Community environment

3.1.10.1 Research institutions

- The R and D centers are considered an influencing factor in determining a petrochemical facility location.

3.1.10.2 Health care system

- The health system of the host country is considered a factor in the study of locating or relocating a petrochemical facility location.

3.1.10.3 Educational system level

- The education system indices used compared with the bench mark educational system around the globe.

3.1.10.4 Business sense

- The community business sense environment.

3.1.10.5 Environmental hazard

- Air pollution, hazard and gases from the production operations to residential areas.

3.1.10.6 Living cost

- This indicator measures the living expenses.

3.1.10.7 Housing availability

- Housing availability and capability of the host country or region where selected for the petrochemical facility to be built at.

3.1.10.8 Climate

- The physical climate of the host country or region where the petrochemical facility located.

3.1.10.9 Religious views

- The religion and culture of the host country also is a measure used in the location decision of a manufacturing facility location.

3.1.10.10 Availability of recreational facilities

- Availability and accessibility to the recreational facility in the host country.

3.1.10.11 Shopping areas

- Malls, shopping centers availability in the host countries.

3.2 World Bank selected factors

In this section, the critical factors of the petrochemical facility location are defined from the World Bank database. These critical factors are the reflection of the defined literature review attributes. The defined critical factors are the backbone of the research. After the data of these factors are collected then the execution of the model is possible.

There are some issues regarding the collection of the data that can be described as first, either the data for the period of time slot are not available or very old, for example, more than five years old. Second, when the data of a certain country for some attributes do not even exist for the entire history, in this case either the country or the attribute is excluded.

The developed model of the research analyzes the set of the critical factors for the chosen countries based on their similarities and dissimilarities, and according to those factors, the countries are assigned to homogenous groups. Each group will have similar attractiveness attributes for locating the petrochemical facility.

3.2.1 GDP

GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in current U.S. dollars.

3.2.2 Lead time to import

Lead time to import is the median time (the value for 50 percent of shipments) from port of discharge to arrival at the consignee. Data are from the Logistics Performance Index survey. Respondents provided separate values for the best case (10 percent of shipments) and the median case (50 percent of shipments). The data are exponentiated averages of the logarithm of single value responses and of midpoint values of range responses for the median case.

3.2.3 Industry, value added

Industry corresponds to international standard industrial classification divisions and includes manufacturing. It comprises value added in mining, manufacturing, construction, electricity, water, and gas. Value added is the net output of a sector after adding up all outputs and subtracting intermediate inputs. It is calculated without making deductions for depreciation of fabricated assets or depletion and degradation of natural resources as a percentage of the GDP.

3.2.4 Inflation, GDP deflator

Inflation as measured by the annual growth rate of the GDP implicit deflator shows the rate of price change in the economy as a whole. The GDP implicit deflator is the ratio of GDP in current local currency to the GDP in constant local currency.

3.2.5 Foreign direct investment, net inflows

Foreign direct investment is the net inflows of investment to acquire a lasting management interest in an enterprise operating in an economy other than that of the investor. It is the sum of equity capital, reinvestment of earnings, other long-term capital, and short-term capital as shown in the balance of payments. This series shows total net foreign direct investment as a percentage of the GDP.

3.2.6 Trade

Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product as a percentage of the GDP.

3.2.7 Lending interest rate

Lending rate is the bank rate 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 financing. The terms and conditions attached to these rates differ by country, however, limiting their comparability.

3.2.8 Services

Services include value added in wholesale and retail trade (including hotels and restaurants), transport, and government, financial, professional, and personal services such as education, health care, and real estate services. Also included are imputed bank service charges, import duties, and any statistical discrepancies noted by national compilers as well as discrepancies arising from rescaling. Value added is the net output of a sector after adding up all outputs and subtracting intermediate inputs. It is calculated without making deductions for depreciation of fabricated assets or depletion and degradation of natural resources

3.2.9 Railways

Goods transported by railway are the volume of goods transported by railway, measured in metric tons times the kilometers traveled.

3.2.10 Air transport

Air freight is the volume of freight, express, and diplomatic bags carried on each flight stage (operation of an aircraft from takeoff to its next landing), measured in metric tons times the kilometers traveled.

3.2.11 Quality of port infrastructure

The Quality of Port Infrastructure measures business executives' perception of their country's port facilities.

3.2.12 Water productivity

Water productivity is calculated as GDP in constant prices divided by the annual total water withdrawal measured in US $ GDP per cubic meter of total freshwater withdrawal.

3.2.13 Cost of business start-up procedures

Cost to register a business is normalized by presenting it as a percentage of gross national income (GNI) per capita.

3.2.14 Unemployment

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

3.2.15 Labor force, total

Total labor force comprises people ages 15 and older who meet the International Labor Organization definition of the economically active population: all people who supply labor for the production of goods and services during a specified period. It includes both the employed and the unemployed. While national practices vary in the treatment of such groups as the armed forces and seasonal or part-time workers, in general the labor force includes the armed forces, the unemployed and first-time job-seekers, but excludes homemakers and other unpaid caregivers and workers in the informal sector.

3.2.16 Research and development expenditure

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. R&D covers basic research, applied research, and experimental development as a percentage of the GDP.

3.2.17 Time required to get electricity

Time required to get electricity is the number of days to obtain a permanent electricity connection. The measure captures the median duration that the electricity utility and experts indicate is necessary in practice, rather than required by law, to complete a procedure.

3.2.18 Trade in services

Trade in services is the sum of service exports and imports divided by the value of GDP, all in current U.S. dollars, as a percentage of the GDP.

3.2.19 Mineral rents

Mineral rents are the difference between the value of production for a stock of minerals at world prices and their total costs of production as a percentage of the GDP.

3.2.20 Oil rents

Oil rents are the difference between the value of crude oil production at world prices and total costs of production as a percentage of the GDP.

3.2.21 Natural gas rents

Natural gas rents are the difference between the value of natural gas production at world prices and total costs of production as a percentage of the GDP.

3.2.22 Total natural resources rents

Total natural resources rents are the sum of oil rents, natural gas rents, coal rents (hard and soft), mineral rents, and forest rents as a percentage of the GDP.

3.2.23 Exports of goods and services

Exports of goods and services represent the value of all goods and other market services provided to the rest of the world. They include the value of merchandise, freight, insurance, transport, travel, royalties, license fees, and other services, such as communication, construction, financial, information, business, personal, and government services as a percentage of the GDP.

3.2.24 Imports of goods and services

Imports of goods and services represent the value of all goods and other market services received from the rest of the world. They include the value of merchandise, freight, insurance, transport, travel, royalties, license fees, and other services, such as communication, construction, financial, information, business, personal, and government services. They exclude compensation of employees and investment income (formerly called factor services) and transfer payments as a percentage of the GDP.

3.2.25 Population

Total population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship. The values shown are midyear estimates.

3.2.26 Manufacturing value added

Manufacturing value added is the net output of a sector after adding up all outputs and subtracting intermediate inputs. It is calculated without making deductions for depreciation of fabricated assets or depletion and degradation of natural resources.

3.2.27 Wage and salaried workers

Wage and salaried workers (employees) are those 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.

3.2.28 Tax income

Taxes on income, profits, and capital gains are levied on the actual or presumptive net income of individuals, on the profits of corporations and enterprises, and on capital gains, whether realized or not, on land, securities, and other assets. Intergovernmental payments are eliminated in consolidation.

CHAPTER FOUR

4.1 Model Description

The main objective of the effective clustering analysis in this research is to cluster the countries into similar groups according to their attractiveness level to a potential petrochemical industrial site. To do so the analysis needs defined critical factors of such an industrial location where the data of these factors are analyzed and implemented in the model of this study.

In Chapter Two, from the extensive literature reviews of the factors affecting and attracting the petrochemical plant location, these critical factors of the petrochemical industrial site have been defined that were frequently quoted in previous research. These publications included a variety of case studies and surveys of the manufacturing sites throughout the globe. Selected factors data were implemented in this research model in quantitative methods.

As mentioned above, the first step in analyzing these factors is to collect the data of the selected factors that have been extracted from the World Bank (WB) database. These data include real numbers of each factor of each country.

4.1.1 Data collection

To obtain the optimal location of a petrochemical plant is to use its critical factors data carefully. These data have to be collected from reliable and trusted sources such as the World Bank data base or any other global data base. These data are highly sensitive and affect the results of the optimal locations.

The local numerical data of the selected attributes can be used to solve the petrochemical facility location problems instead of the major indices that are listed in the globally recognized databases. In this research, the numerical data have been collected from the World Bank data base because it has many advantages over the other indices' data base. First, the WB data base has more distinctive factors. Second, the WB data base not only has more factors but also has yearly updated indices. Third, it is the authentic reliable source for the global indices in general that make it the number one source for decision makers around the globe.

The sensitivity of collecting the data in some countries due to their own political or interest status makes this mission costly and complicated. As a result, there is a lack of enough information about some countries that are considered an important and potential location for petrochemical plants. There are some issues regarding the collected data from the WB data base as well. The challenges of collecting the data are in two categories. The first category is where the selected factors from the literature cannot be found as is. In this case the factor could be sub factored as an example: *Transportation factors could be sub factored to:*

- *Number of international airports*
- *Number of sea ports to the country*
- *Railways throughout the country*
- *Highways system of the country*
- *Infrastructure of the country*

So, when a certain factor has been sub factored and the data of these sub factors have been collected, all are counted towards the main factor to assure higher accuracy. The second category is when the factor data of some countries cannot be found in the WB data base and cannot be sub factored as well. This case can be solved in two ways: first, by looking to the missing factor data of the current year, several slots of the previous years can be used by creating a trend used to forecast the data for the needed time slot or by looking up different global data bases. Second, if the data of a certain factor are missing for the time periods, this factor will be eliminated. In the proposed model, the collected data of the selected critical factors from the WB data base are expressed in different ranges of values, some in decimals and some in billions. Thus, before the analysis is applied and the countries are clustered into groups, the first step after collecting the data from the WB data base is normalizing the collected data by the formula:

$$X_{normalized} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

In this formula, all the collected data numerical values will land in a range between 0-1.

4.1.2 Model implementation

A hierarchical clustering algorithm applied in this research model approach starts by analyzing a singular object one by one until it forms similar behavior objects in groups according to their attractiveness level to a petrochemical facility location. Similar objects are assigned in a homogenous group at the end of the algorithm implementation. Moreover, the number of the clusters can be defined and the objects accordingly will be assigned differently into their groups because the number of clusters is proportional to a closer behavior of the objects (countries) towards the petrochemical facility location.

The clustering algorithm consists of three main aspects:

- Objects
- Attributes
- Similarity coefficient

4.2.2.1 Objects

Objects are the countries in this research clustering algorithm model. This clustering technique applied to gather the similar countries together in groups is based in their similarities and dissimilarities in terms of the selected factors in the model. Thus, the similar objects are assigned together in one group and the dissimilar are assigned in a different group.

4.2.2.2 Attributes

Attributes are the selected critical factors upon which the clustering analysis is applied. In this research, the attributes are the primary critical and sensitive part of the research because the cluster formations are directly affected according to the selected factors. Any misleading data of any factor will result in assigning the country to nonhomogeneous clusters. As defined in Chapter Three, the critical factors of this research can be analyzed to form the clusters based on their similarities and dissimilarities by considering all factors together.

4.2.2.3 Similarity coefficient

Euclidean distance is the similarity coefficient used for measuring data on the same scale. This similarity coefficient formula is:

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \qquad (2)$$

where x and y are two vectors and are often used to compare profiles of respondents across variables.

4.1.3 Clustering method

The method implemented in this research is the complete-linkage that is one of the agglomerative hierarchical clustering methods. There are several methods of clustering analysis that combine the similar objects together in a cluster by measuring the shortest distance between the two objects. The "shortest distance" is what makes the differences between the clustering methods. In complete-linkage clustering (CLINK) the relation between two clusters involves all the pair elements inside the two tested clusters so it applies for the farthest distance between the two elements in each analyzed cluster. After these distances are measured, the CLINK starts forming the clusters according to the distance between the pairs of elements in each cluster.

$$d_{complete}(r,s) = \underset{i \in r, j \in s}{Max}\, d_{ij} \tag{3}$$

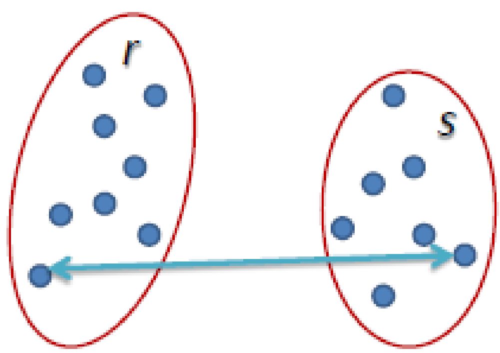

Figure 1: Complete linkage measuring distance method

CLINK measures the longest distance between a single element to another single element in each cluster and then combines the clusters based on the shortest distances of this measurement method. CLINK has advantages over the other methods because of its own characteristics:

- CLINK prevents the merger of two clusters together for only a high level of similarity between two members while other members are dissimilar (chaining reaction problem of some clustering algorithms)
- The least similar pair between two clusters is used to determine the inter-cluster similarity
- The clusters are small and tightly bound
- CLINK is computer software-friendly (MATLAB).

Average linkage clustering is one of the hierarchal agglomerative technique to form the clusters based on the calculated distance between the clusters. It measures the average distance between the elements of the two clusters to determine the distance between the respective clusters as shown in equation 4.

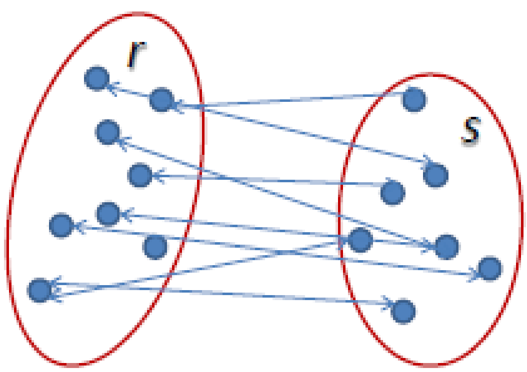

Figure 2: Average linkage measuring distance method

$$L(r,s) = \frac{1}{n_r n_s} \sum_{i=1}^{n_r} \sum_{j=1}^{n_s} D(x_{ri}, x_{sj}) \quad (4)$$

The city block distance is very similar to the Euclidean distance but it uses the following equation to measure the distance between two data points.

$$\sum_{i=1}^{K} |a_i - b_i| \quad (5)$$

4.1.4 Results and dendrograms

When the numerical data of the selected critical factors are collected for the targeted countries in the model, then the model is ready to be analyzed. Clustering analysis can be applied to the numerical values in the model by using the MATLAB software codes. MATLAB has enabled hierarchical techniques in order to form the clusters of the similar countries in groups.

To implement the selected method of the clustering analysis to the gathered numerical values, the following steps need to be done before the clustering can take place:

- All the factors and countries need to be listed with their numerical values without missing a slot in an Excel spread sheet.
- Normalization has to be done to the numerical values due to the huge difference between the numerical values to convert those in billions and those in decimals all to a range of 0-1.
- The MATLAB codes need to be prepared for the selected method.

In MATLAB operations each function has its code and the following illustrates what code does what:

❖ *Pdist (A)* is the function built in the MATLAB to calculate by default the Euclidean distance between the countries.

❖ *tree=linkage(D,'complete');*

[~, T]=dendrogram(tree,x)

After the clusters are assigned to each country according to the similarity coefficient by using the Euclidean distance method, this code function shows these clusters in diagrams called *dendrograms.*

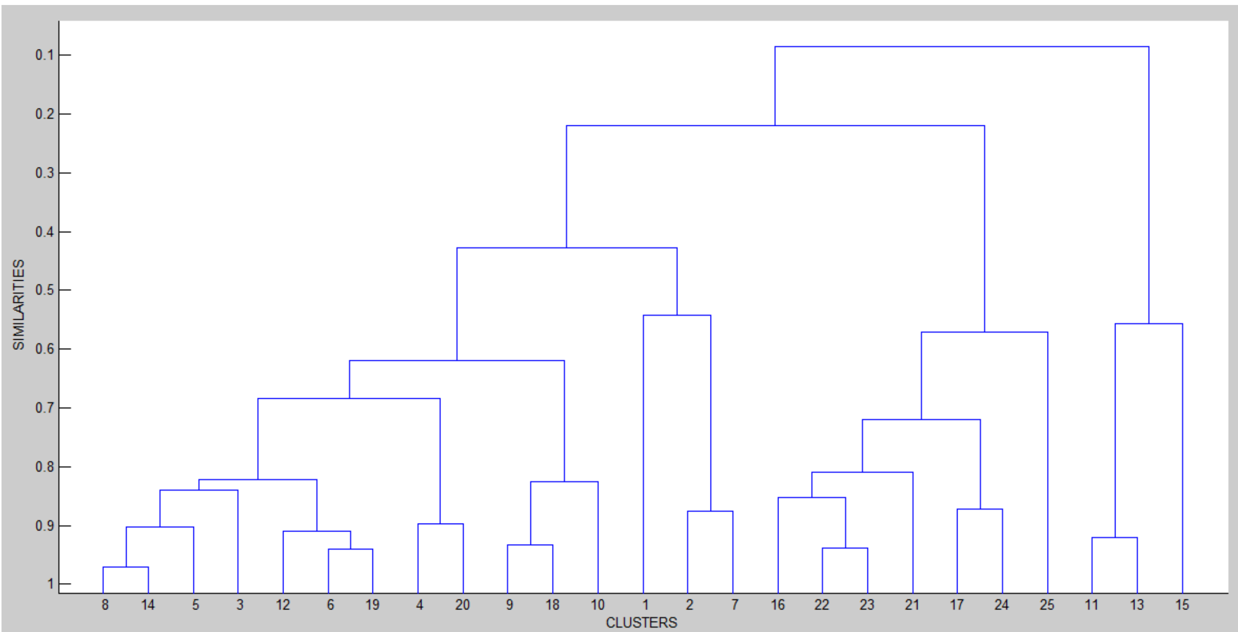

Figure 3: Dendogram shows the similarity in y-axis and clusters in x-axis

Euclidean distance is another similarity coefficient used for measuring data on the same scale. This similarity coefficient formula is

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \quad (6)$$

where x and y are two vectors and it is often used to compare profiles of respondents across variables.

4.1.5 Robustness of the model

This thesis of developing a model that quantifies the factors towards their corresponding locations gives a new horizon to comprehensive clustering analysis. The clustering analysis of using complete-linkage (CLINK) assign the countries in groups according to their attractiveness level to the petrochemical facility location. These groups are homogenous within the group and also differ from one another. The model has real numerical values from the World Bank database so with using this kind of number that does have a trend in most cases; the robustness is going to be demonstrated by the following steps:

1. The CLINK cluster will be applied to the model numerical value of 30 countries for four stages as follow:
    - Considering only 3 factors.
    - Considering up to 6 factors.
    - Considering up to 12 factors.

If the countries' clusters keep changing with adding more factors, then the model numerical factors are making a difference and the clustering analysis takes them in account.

2. Apply a different similarity coefficients method to demonstrate the robustness of the results collected from the similarity coefficient with the Euclidean distance method.

4.2 Model challenges in data collection

The challenges of collecting the data can be in two categories. The first category is where the selected factors from the literature cannot be found as is. In this case the factor can be sub factored . When a certain factor has been sub factored and the data of these sub factors have been collected they are all counted towards the main factor to assure higher accuracy. The second category is when the factor data of some countries cannot be found in the WB data base and cannot be sub factored as well. This case can be solved by two ways: first, by looking at the missing factor data of the several slots of the previous years or by looking up a different global data base. Second, if the data of a certain factor are missing for the time periods this factor will be eliminated.

4.3 Locations ranking

The weight of the ranking will be assigned to each factor. Giving the selected critical factors a weight in the developed model according to how strong is each impact to the petrochemical facility location will show a closer analysis to the clusters formed.

The previous research, surveys and case studies repeatedly emphasize certain factors more than others. Some studies listed the factors in a sequence of importance, with first being the most important factor.

Factor weight is ranking the countries first; the highest is the most potential location to the petrochemical facility; ranking the countries within their own cluster will give more emphasis to the petrochemical industry decision makers and investors.

Table 1: Weight assigned to the critical attributes

Developed model selected critical factors		Factor weight
1	Cost of business start-up procedures	0.04
2	Time required to start a business	0.04
3	Time required to get electricity (days)	0.04
4	Air transport, freight (million ton-km)	0.01
5	Lead time to import	0.01

6	Industry, value added (annual % growth)	0.04
7	Manufacturing, value added (% of GDP)	0.04
8	Industry, value added	0.04
9	Exports of goods and services	0.04
10	Imports of goods and services	0.04
11	Quality of port infrastructure	0.04
12	Foreign direct investment	0.06
13	Labor force	0.05
14	Wage and salaried workers	0.05
15	Unemployment %	0.04
16	Internet Users	0.02
17	GDP	0.08
18	Trade (% of GDP)	0.03
19	Trade in services (% of GDP)	0.04
20	Inflation, GDP deflator	0.06
21	Tax on income	0.03
22	Tax revenue	0.03
23	Mineral rents	0.02
24	Oil rents	0.02
25	Natural gas rents	0.02
26	Lending interest rate	0.01
27	Services	0.04
28	Population	0.02

The following is an example of how the complete linkage clustering calculates the distance between a pair of elements needed to be clustered. The distance between the pair of elements would be symmetric. The distance between the country here and itself is zero as shown in the table below. The table has identical values around the diagonal.

Table 2: The measured distance between the objects

	Country 1	**Country 2**	**Country 3**	**Country 4**	**Country 5**
Country 1	0	9	3	6	11
Country 2	9	0	7	5	10
Country 3	3	7	0	9	2
Country 4	6	5	9	0	8
Country 5	11	10	2	8	0

Table 3: The measured distance between the objects diagonal

	Country 1	**Country 2**	**Country 3**	**Country 4**	**Country 5**
Country 1	0				
Country 2	9	0			
Country 3	3	7	0		
Country 4	6	5	9	0	
Country 5	11	10	2	8	0

Hence one triangle used for the measured distance between the clusters. As indicated in the previous table, the smallest value is the country three to the country five so they merged up a cluster country three and five. Since the complete linkage the maximum value would be used for the forming the new matrix of the distances. For the distance (1, 3) = 3 and the distance (1, 5) = 11 so the value 11 should be selected in the new measure matrix, as shown in table 1 and 2.

Table 4: The first iteration of the measured distance between the objects diagonal

	Country 3 &5	Country 1	Country 2	Country 4
Country 3&5	0			
Country 1	11	0		
Country 2	10	9	0	
Country 4	9	6	5	0

The next shortest distance is between country 2 and country 4. Then the same iterations process continues to the final iteration which shapes this dendogram, as shown in Figure 4.

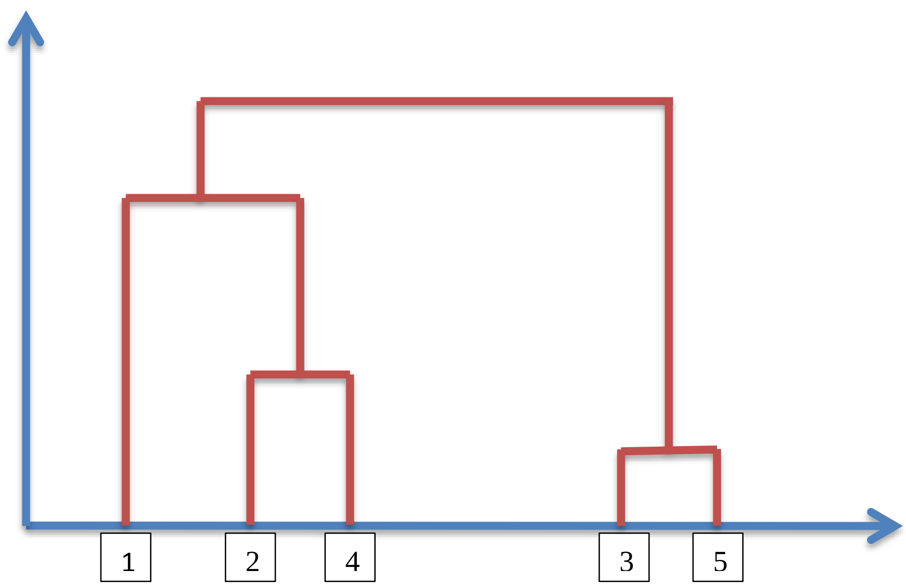

Figure 4: Dendogram of the assigned clusters

The dendogram above shows the closer distance between the countries assigned together in the same cluster. The y-axis shows the distance between the clustered objects.

CHAPTER FIVE

Developed Model Results and Discussion

5.1 Developed Model

In this chapter, the collected data of the developed model and the selected critical factors have been finalized. The clustering analysis of the objects (countries) of the developed model will take place to propose the optimum set of locations for the petrochemical facility. The implemented similarity coefficient method of this study and the clustering analysis is complete-a linkage clustering algorithm with Euclidean distance.

The validity of the developed model can be tested by applying the complete-linkage clustering analysis (CLINK) method to different sizes of numerical values of 30 countries for three stages as follows:

- Execute the model with only 3 factors.
- Execute the model with only 6 factors.
- Execute the model with only 12 factors.

This test proves the functionality and flexibility of the proposed model by allowing this unique feature to compare the facility location problem solving methods where the factors or the countries can be added or eliminated. That change has its own impact on the clustering results afterwards. Then the bigger numerical values of the developed model, 100 countries with 28 factors each, will be subjected to the same clustering analysis test that is complete-linkage clustering by using the Euclidean distance method.

For the robustness of the results that have been generated from the previous method of clustering, the last approach of this study will include the implementation of different clustering methods:

- Implementation of Euclidean distance with complete-linkage clustering
- Implementation of CityBlock with complete-linkage clustering
- Implementation of Euclidean distance with average-linkage clustering
- Implementation of CityBlock with average-linkage clustering

5.2 Top 30 Countries Analysis

In this section, as essential step of applying the complete-linkage clustering analysis to the developed model is to select the list of objects (countries) where they can be tested and clustered based on their similarities and dissimilarities of the selected attributes. Then the homogenous clusters of countries having a similar behavior towards the petrochemical facility location will be measured and analyzed.

As one step of validating this developed model, the beginning will be with the top 30 GDP countries where they are subjected to multiple sessions of multiple factors. The complete-linkage clustering will be used for any analysis and dendrograms unless other methods are declared. The GDP is the major indicator of a nation's economy. The GDP measures multiple aspects of the nation:

- Buying power
- Economy Size
- Long and short term reliability of economy measures
- Standard of living

All of these data were collected from the most reliable and recognized global data base, the World Bank database. The list of the top 30 GDP nations are listed below and the list is ranked from 1 to 30 with the highest GDP listed first.

Table 5: The list of the top 30 GDP nations

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

19	Switzerland
20	Saudi Arabia
21	Argentina
22	Sweden
23	Nigeria
24	Poland
25	Belgium
26	Thailand
27	Norway
28	Austria
29	United Arab Emirates
30	Egypt, Arab Republic

5.2.1 The model selected critical factors

After the first step of listing the selected countries for the developed model, the next step to execute the clustering analysis is to decide which critical factors are considered in this model. The numerical values of these factors were collected from the World Bank database. Because of missing or unavailable data for some factors, the more comprehensive factors are considered, for example, the GDP, when it is implicitly covers more than one factor. The factors directly related to the petrochemical industry are considered confidential information and put more challenges on the data collection. The factors considered in this developed model are listed below:

Table 6: Developed model selected critical factors

Developed model selected critical factors

1	Cost of business start-up procedures	15	Unemployment %
2	Time required to start a business	16	Internet Users
3	Time required to get electricity (days)	17	GDP
4	Air transport, freight (million ton-km)	18	Trade (% of GDP)
5	Lead time to import	19	Trade in services (% of GDP)
6	Industry, value added (annual % growth)	20	Inflation, GDP deflator
7	Manufacturing, value added (% of GDP)	21	Tax on income
8	Industry, value added	22	Tax revenue
9	Exports of goods and services	23	Mineral rents
10	Imports of goods and services	24	Oil rents
11	Quality of port infrastructure	25	Natural gas rents
12	Foreign direct investment	26	Lending interest rate
13	Labor force	27	Services
14	Wage and salaried workers	28	Population

5.2.2 Model collected data

The model data were collected from the World Bank database that represents the numerical values of the selected critical factors of the study. The data of this model are available in the appendix.

5.2.3 Applied clustering technique

In this step, the developed model has been set up with the collected data and is ready to implement the clustering analysis method. Implementing the complete-linkage clustering

(CLINK) with the Euclidean distance coefficient to the data will result in multiple clusters of countries that have similar behavior towards the selected factors.

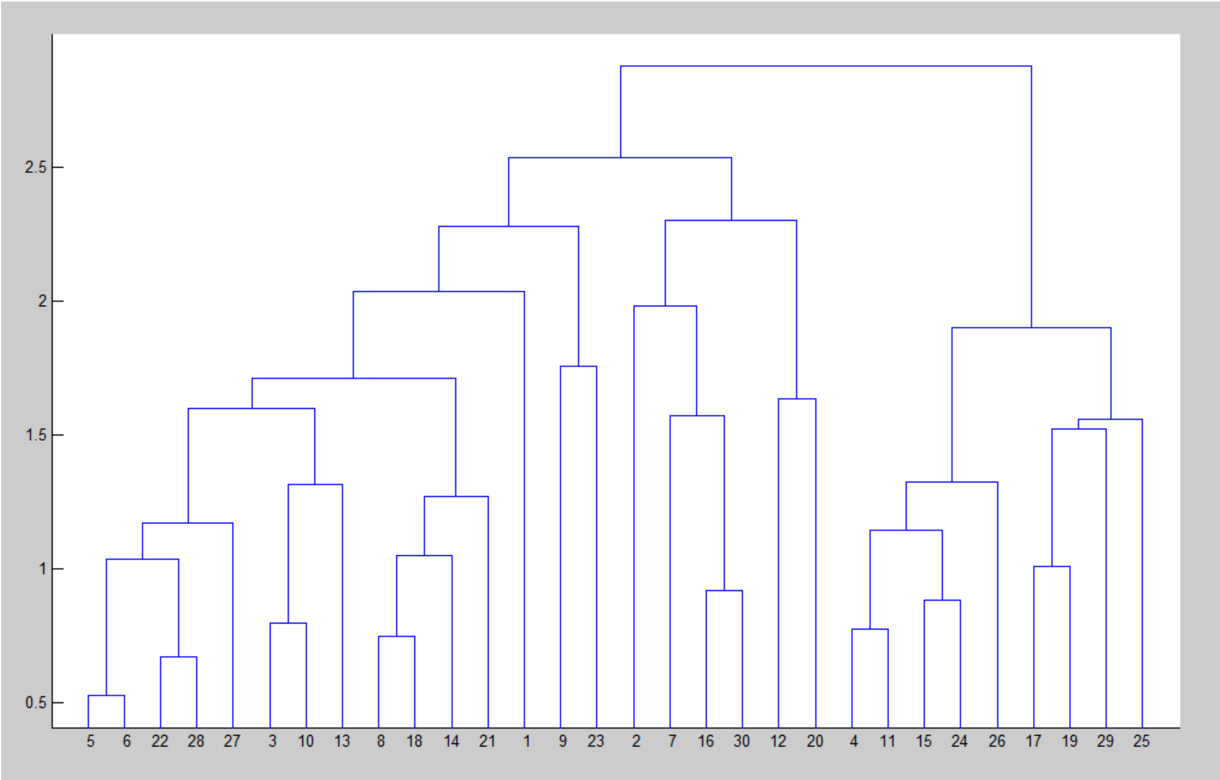

Figure 5: 30 Clusters of all factors dendrogram of the complete-linkage by Euclidean coefficient Implemented.

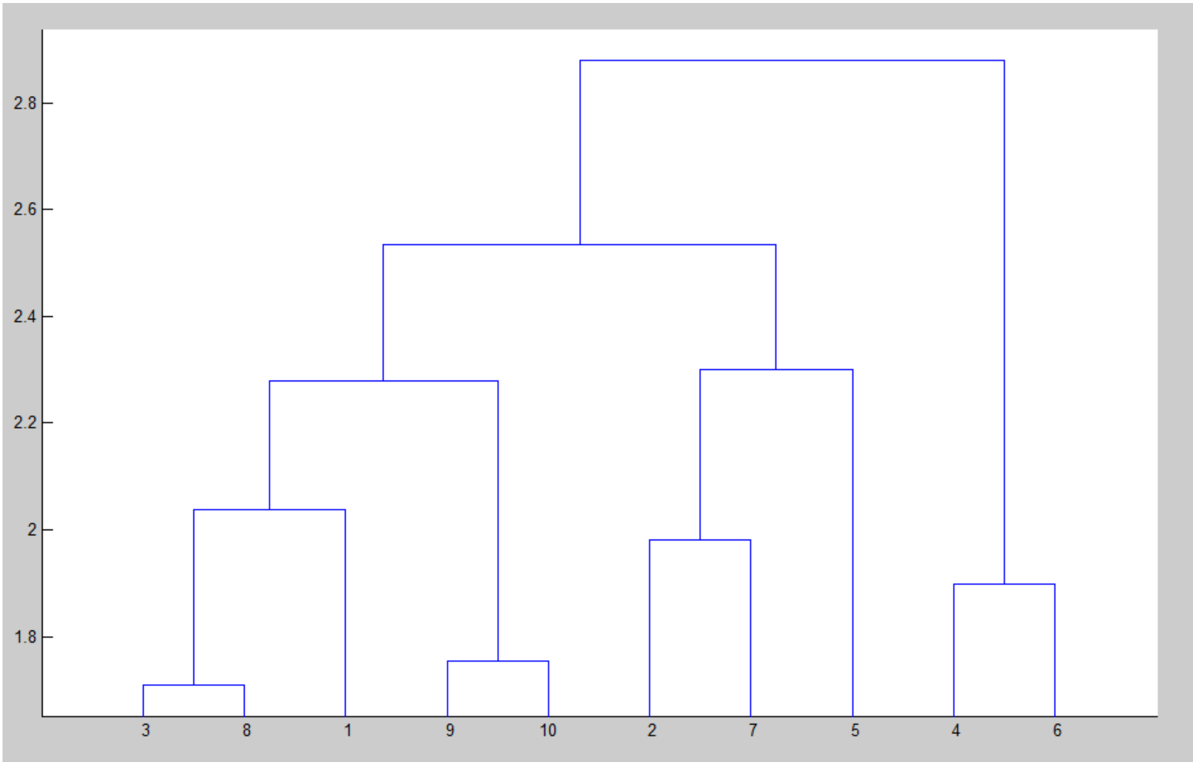

Figure 6: 10 Clusters of all factors dendrogram of the complete-linkage by Euclidean coefficient Implemented.

Table 7: The top 30 countries were grouped into 10 clusters.

Country	Cluster Number
United States	1
China	2
Japan	3
United Kingdom	3
France	3
Canada	3
Australia	3

Sweden	3
Norway	3
Austria	3
Germany	4
Korea, Republic of	4
Mexico	4
Poland	4
Thailand	4
Russian Federation	5
Saudi Arabia	5
Netherlands	6
Switzerland	6
Belgium	6
United Arab Emirates	6
India	7
Indonesia	7
Egypt, Arab Republic	7
Italy	8
Spain	8
Turkey	8
Argentina	8
Brazil	9
Nigeria	10

5.2.4 Ranking and factors weight

The clusters are formed according to the selected critical factors from the collected data. Some critical factors directly affect the global petrochemical facility location more than others. The clusters are formed based on the behavior of the country towards the petrochemical plant site by analyzing the impact of the entire selected factors to give clusters of countries that have the same attractiveness level to the petrochemical facility location.

Providing the selected critical factors in the developed model according to the strength of the impact of each on the petrochemical facility location will show a closer analysis to the clusters formed. Ranking the highest whole locations (countries) for the most potential location for the petrochemical facility location as well as ranking the countries within their own cluster will give more information to the petrochemical industry decision makers and investors.

The weight of the ranking will be assigned to each factor based on the impact of the factor to a potential petrochemical facility location. The previous research, surveys and case studies repeatedly emphasize certain factors more than others. Some studies listed the factors in a sequence of importance:

Table 8: The general ranking of the top 30 countries

Cluster #	Country Name	Ranking of the 30 countries
1	United States	1
2	China	2
	Japan	3
	United Kingdom	5
	France	6

3	Canada	10
	Australia	13
	Sweden	22
	Norway	27
	Austria	28
4	Germany	4
	Korea, Republic	11
	Mexico	15
	Poland	24
	Thailand	26
5	Russian Federation	12
	Saudi Arabia	20
6	Netherlands	17
	Switzerland	19
	Belgium	25
	United Arab Emirates	29
7	India	7
	Indonesia	16
	Egypt, Arab Republic	30
8	Italy	8
	Spain	14
	Turkey	18
	Argentina	21

9	Brazil	9
10	Nigeria	23

Also, a ranking within the cluster could be done easily after the countries are categorized in whole countries ranking in the previous table. The next table will show each cluster group of countries ranked within the cluster, from highest to lowest.

Table 9: The ranking of the countries within the cluster

Cluster #	Country Name	Ranking within the cluster
1	United States	1
2	China	1
3	Japan	1
	United Kingdom	2
	France	3
	Canada	4
	Australia	5
	Sweden	6
	Norway	7
	Austria	8
4	Germany	1
	Korea, Republic	2
	Mexico	3
	Poland	4

	Thailand	5
5	Russian Federation	1
	Saudi Arabia	2
6	Netherlands	1
	Switzerland	2
	Belgium	3
	United Arab Emirates	4
7	India	1
	Indonesia	2
	Egypt, Arab Republic	3
8	Italy	1
	Spain	2
	Turkey	3
	Argentina	4
9	Brazil	1
10	Nigeria	1

5.3 Impact of the number of critical factors considered

One of the major advantages of this developed model is its own flexibility and sensitivity of adding or eliminating objects or attributes. In addition, the model will give stronger bond relations proportional to the factors added to it. To measure that effect of adding more factors would change the clusters of the countries. This will take place for the 30 top GDP countries with three setups. The first is applying three factors, then six factors and then the last twelve factors.

To apply these steps consistently, the number of clusters would be ten clusters for each clustering analysis of each set of factors.

5.3.1 The developed model analysis with three factors considered

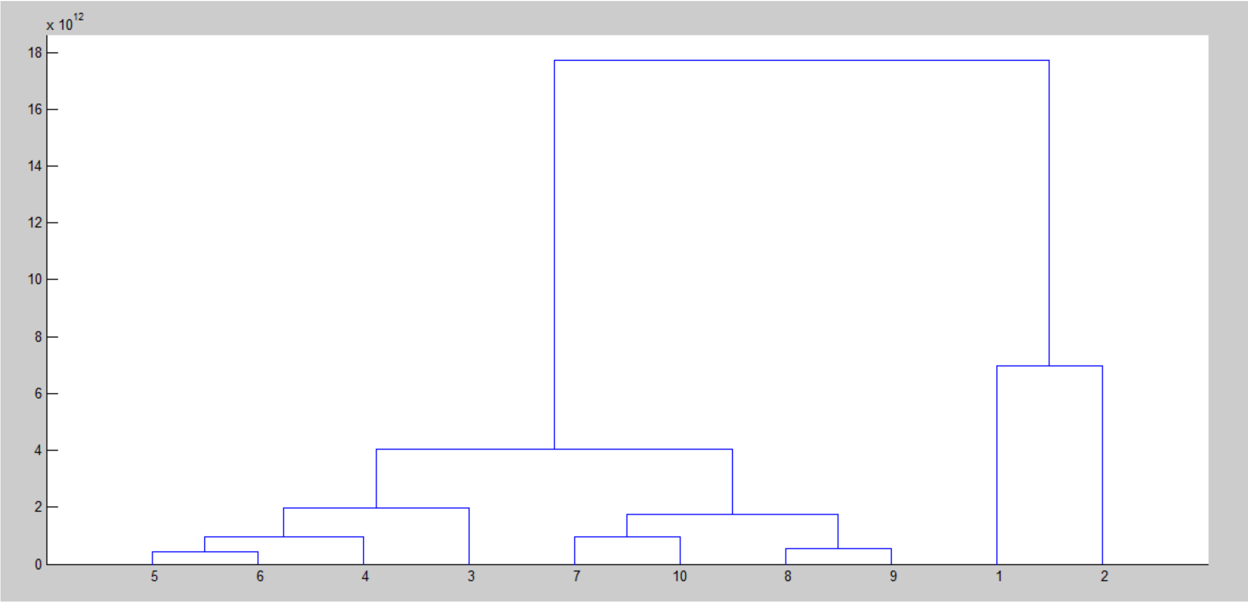

Figure 7: Complete-linkage clustering with Euclidean distance applied to the top 30 countries in 10 clusters considering three factors

Table 10: Complete-linkage clustering with Euclidean distance applied to the top 30 countries in 10 clusters considering three factors

Country	Cluster Number
United States	1
China	2
Japan	3
Germany	4
United Kingdom	5
France	6
India	7
Italy	7
Brazil	7
Canada	10
Korea, Republic	10
Russian Federation	10
Australia	10
Spain	10
Mexico	10
Indonesia	8
Netherlands	8
Turkey	8
Switzerland	8
Saudi Arabia	8
Argentina	8

Sweden	9
Nigeria	9
Poland	9
Belgium	9
Thailand	9
Norway	9
Austria	9
United Arab Emirates	9
Egypt, Arab Republic	9

5.3.2 The developed model analysis with six factors considered

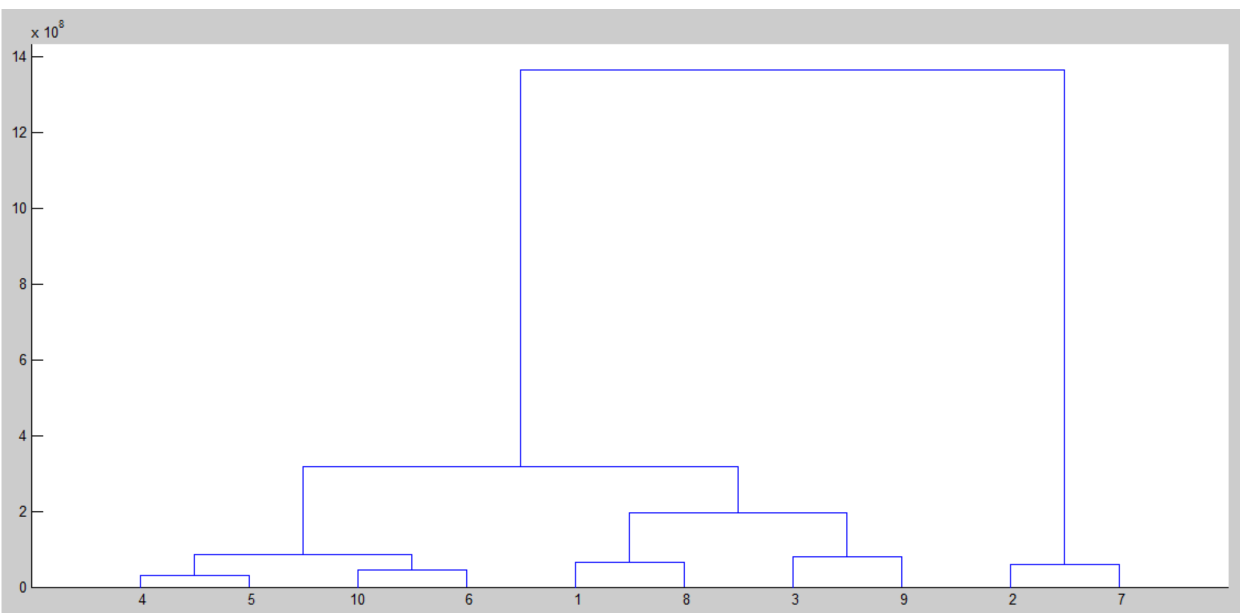

Figure 8: Complete-linkage clustering with Euclidean distance applied to the top 30 countries in 10 clusters considering six factors

Table 11: Complete-linkage clustering with Euclidean distance applied to the top 30 countries in 10 clusters considering six factors

Country	Cluster Number
United States	1
China	2
Japan	3
Germany	4
United Kingdom	5
France	5
India	7
Italy	5
Brazil	9
Canada	10
Korea, Republic	10
Russian Federation	3
Australia	6
Spain	10
Mexico	3
Indonesia	8
Netherlands	6
Turkey	4
Switzerland	6
Saudi Arabia	10
Argentina	10

Sweden	6
Nigeria	9
Poland	10
Belgium	6
Thailand	5
Norway	6
Austria	6
United Arab Emirates	6
Egypt, Arab Republic	4

5.3.3 The developed model analysis with 12 factors considered

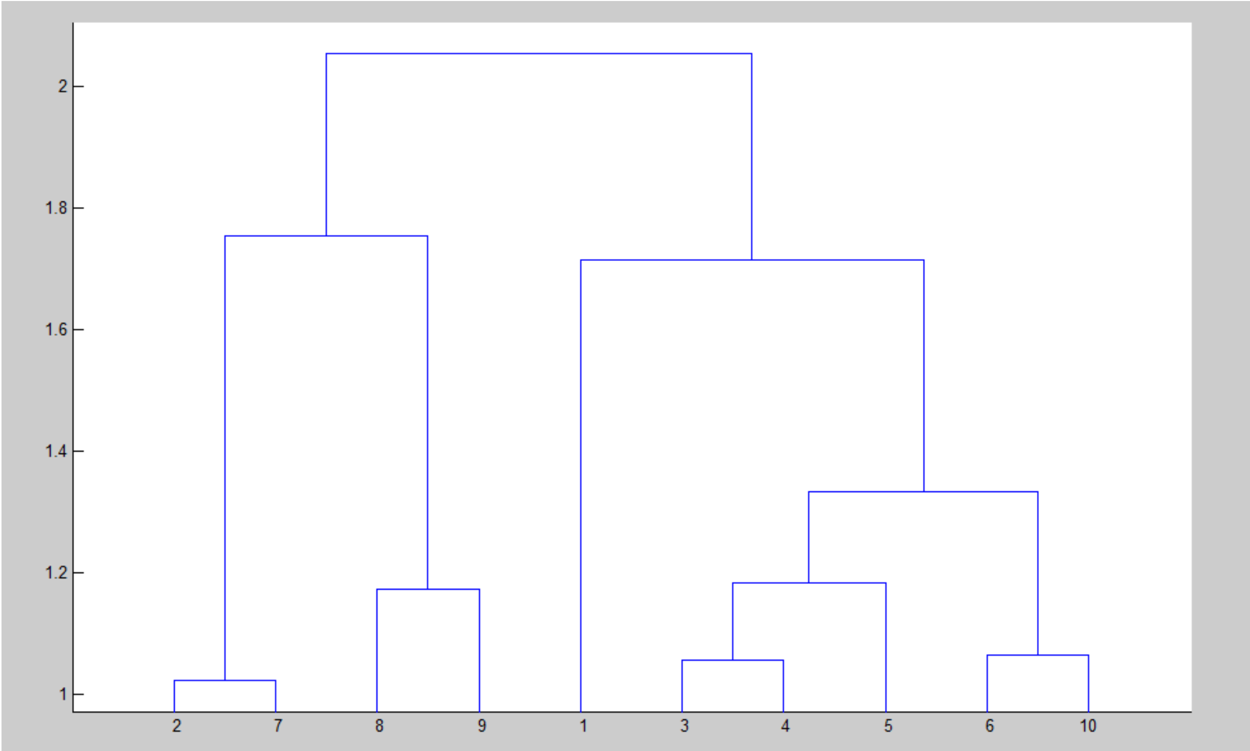

Figure 9: Complete-linkage clustering with Euclidean distance applied to the top 30 countries in 10 clusters considering 12 factors

Table 12: Complete-linkage clustering with Euclidean distance applied to the top 30 countries in 10 clusters considering 12 factors

Country	Cluster Number
United States	1
China	2
Japan	3
Germany	3
United Kingdom	3
France	3
India	7
Italy	8
Brazil	9
Canada	3
Korea, Republic	3
Russian Federation	8
Australia	3
Spain	4
Mexico	8
Indonesia	8
Netherlands	5
Turkey	8
Switzerland	5
Saudi Arabia	3
Argentina	8

Sweden	3
Nigeria	8
Poland	8
Belgium	6
Thailand	10
Norway	3
Austria	3
United Arab Emirates	6
Egypt, Arab Republic	8

In the table below the cluster results of the set of factors for each country are compared and the clusters change as more factors are added to the developed model. Because the model is flexible, it is most likely proportional for the number of factors added to it. According to the results of the clustering analysis of the different sets of factors, the developed model of the research will have comprehensive clusters because it can execute a higher number of objects and attributes than other models.

Table 13: Functionality of the model by the comparison of adding set of factors.

Country	**3 Factors**	**6 Factors**	**12 Factors**
United States	1	1	1
China	2	2	2
Japan	3	3	3
Germany	4	4	3
United Kingdom	5	5	3

France	6	5	3
India	7	7	7
Italy	7	5	8
Brazil	7	9	9
Canada	10	10	3
Korea, Republic	10	10	3
Russian Federation	10	3	8
Australia	10	6	3
Spain	10	10	4
Mexico	10	3	8
Indonesia	8	8	8
Netherlands	8	6	5
Turkey	8	4	8
Switzerland	8	6	5
Saudi Arabia	8	10	3
Argentina	8	10	8
Sweden	9	6	3
Nigeria	9	9	8
Poland	9	10	8
Belgium	9	6	6
Thailand	9	5	10
Norway	9	6	3
Austria	9	6	3

United Arab Emirates	9	6	6
Egypt, Arab Republic	9	4	8

5.4 Impact of the weights assigned to the model critical factors

The clusters are formed according to the selected critical factors and the collected data. There are some critical factors that directly affect the global petrochemical facility location more than the others. The clusters form based on the behavior of the country towards the petrochemical plant site by analyzing the impact of the entire selected factors to provide clusters of countries that have the same attractiveness level for the petrochemical facility location.

Giving the selected critical factors in the developed model according to how strong an impact each has to the petrochemical facility location will show a closer analysis of the clusters formed. Ranking the whole locations (countries) the highest as the most potential location for the petrochemical facility location as well as ranking the countries within their own cluster will give more information to the petrochemical industry decision makers and investors.

As mentioned earlier, the weight of the ranking will be assigned to each factor based on the impact of the factor on a potential petrochemical facility location. The previous research, surveys and case studies repeatedly emphasize certain factors more than others. Some studies listed the factors in a sequence of importance, listing the most important factor first.

The weight of the factors used in the analysis of the top 30 GDP countries will be applied for the bigger model of this research. The targeted factors will be weighed higher than the others with the ability of having different sets of weighted factors.

5.5 The model developed from this research

The sample consists of 100 countries that will cluster in homogenous groups according to their attractiveness level to a potential petrochemical facility location. In this model, the flexibility is one of the major components and it can be described by three points:

- The model can function with a small sample of data or with big samples.
- The model is flexible in adding or eliminating attributes.
- The model has the ability to analyze real time data.

For more robustness, another similarity coefficient approach was applicable to be used to test the real-time data and compare them with the research clusters results that used complete-linkage clustering. Several approaches of similarity coefficients with different clustering algorithm will be implemented.

5.5.1 Defined a list of countries

A real-world sample of countries will be analyzed and clustered. The sample is a hundred countries of the most attractive locations for a potential petrochemical facility location. Those countries have the highest GDP. The numerical values of these countries along with their factors will shape up the clustering analysis that measures the similarities and dissimilarities among them. In this table below the countries of the study are listed in a sequence with highest GDP first:

Table 14: the list of the countries of the research sample

#	**Country**	#	**Country**	#	**Country**
1	United States	34	Malaysia	67	Kenya
2	China	35	Singapore	68	Myanmar

3	Japan	36	Philippines	69	Ethiopia
4	Germany	37	Colombia	70	Luxembourg
5	United Kingdom	38	Ireland	71	Belarus
6	France	39	Pakistan	72	Costa Rica
7	India	40	Chile	73	Uruguay
8	Italy	41	Finland	74	Azerbaijan
9	Brazil	42	Portugal	75	Panama
10	Canada	43	Bangladesh	76	Bulgaria
11	Korea, Republic	44	Greece	77	Croatia
12	Russian Federation	45	Vietnam	78	Lebanon
13	Australia	46	Peru	79	Tanzania
14	Spain	47	Czech Republic	80	Tunisia
15	Mexico	48	Kazakhstan	81	Slovenia
16	Indonesia	49	Iraq	82	Lithuania
17	Netherlands	50	Romania	83	Yemen, Republic
18	Turkey	51	New Zealand	84	Ghana
19	Switzerland	52	Algeria	85	Jordan
20	Saudi Arabia	53	Qatar	86	Serbia
21	Argentina	54	Hungary	87	Congo
22	Sweden	55	Kuwait	88	Uganda
23	Nigeria	56	Angola	89	Paraguay
24	Poland	57	Morocco	90	Latvia
25	Belgium	58	Ecuador	91	El Salvador

26	Thailand	59	Sudan	92	Trinidad and Tobago
27	Norway	60	Ukraine	93	Estonia
28	Austria	61	Slovak Republic	94	Nepal
29	United Arab Emirates	62	Sri Lanka	95	Zambia
30	Egypt, Arab Republic	63	Oman	96	Honduras
31	South Africa	64	Dominican Republic	97	Cyprus
32	Hong Kong	65	Uzbekistan	98	Afghanistan
33	Denmark	66	Guatemala	99	Cambodia
100	Iceland				

5.5.2 The model selected critical factors

The selected factors of the research as listed before are based on their numerical values and the clustering analysis will assign the countries to groups.

Table 15: The list of the developed factors of the research

Developed model selected critical factors			
1	Cost of business start-up procedures	15	Unemployment %
2	Time required to start a business	16	Internet Users
3	Time required to get electricity (days)	17	GDP
4	Air transport, freight (million ton-km)	18	Trade (% of GDP)
5	Lead time to import	19	Trade in services (% of GDP)
6	Industry, value added (annual % growth)	20	Inflation, GDP deflator
7	Manufacturing, value added (% of GDP)	21	Tax on income

8	Industry, value added	22	Tax revenue
9	Exports of goods and services	23	Mineral rents
10	Imports of goods and services	24	Oil rents
11	Quality of port infrastructure	25	Natural gas rents
12	Foreign direct investment	26	Lending interest rate
13	Labor force	27	Services
14	Wage and salaried workers	28	Population

5.5.3 Model collected data

The data of the hundred countries and critical factors of the study were collected from the World Bank database. The model, along with the numerical values of each country and each factor, is available in the appendix.

5.5.4 Applied clustering technique

The real-world sample to be tested by the complete-linkage clustering (CLINK) method with the Euclidean distance coefficient to the data will result in multiple clusters of countries that have similar behavior based on the selected factors.

5.5.5 Ranking and factors weight

The weight of the ranking will be assigned to each factor based on the impact of the factor on a potential petrochemical facility location. The previous research, surveys and case studies repeatedly emphasize certain factors more than others. Some studies listed the factors in a sequence of importance, with the most important factor listed first.

Table 16: The weight assigned to the critical factors of the research

	Developed model selected critical factors	Factor weight
1	Cost of business start-up procedures	0.04
2	Time required to start a business	0.04
3	Time required to get electricity (days)	0.04
4	Air transport, freight (million ton-km)	0.01
5	Lead time to import	0.01
6	Industry, value added (annual % growth)	0.04
7	Manufacturing, value added (% of GDP)	0.04
8	Industry, value added	0.04
9	Exports of goods and services	0.04
10	Imports of goods and services	0.04
11	Quality of port infrastructure	0.04
12	Foreign direct investment	0.06
13	Labor force	0.05
14	Wage and salaried workers	0.05
15	Unemployment %	0.04
16	Internet Users	0.02
17	GDP	0.08
18	Trade (% of GDP)	0.03
19	Trade in services (% of GDP)	0.04
20	Inflation, GDP deflator	0.06
21	Tax on income	0.03

22	Tax revenue	0.03
23	Mineral rents	0.02
24	Oil rents	0.02
25	Natural gas rents	0.02
26	Lending interest rate	0.01
27	Services	0.04
28	Population	0.02

5.5.6 Complete-linkage clustering of the developed model

In this section, the complete-linkage clustering method with Euclidean distance are employed to analyze the developed model of the research and to measure the similarities and dissimilarities between the objects of the model (countries). As the previous results are collected and the dendrograms formed by using the MATLAB, the same method will be implemented in the real-world sample to form the distinct clusters of the countries.

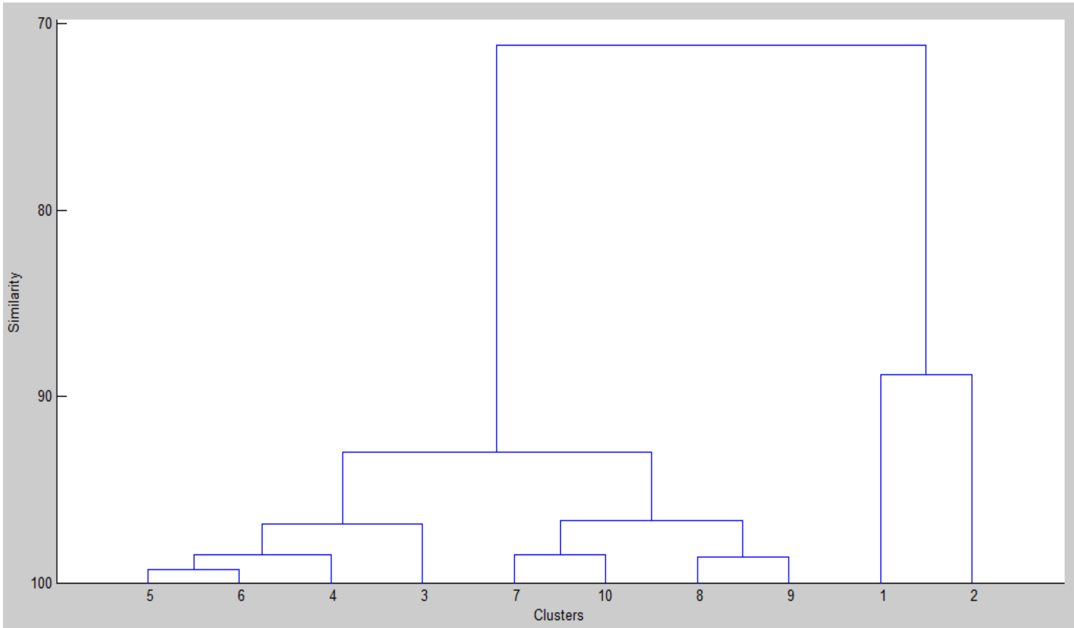

Figure 10: The distinct 10 clusters of the real-world sample of 100 countries by complete-linkage clustering with Euclidean distance

Table 17: The distinct clusters of the real-world sample of 100 countries by complete-linkage clustering with Euclidean distance for 10 clusters and 25 clusters

Country	10 Clusters	25 Clusters	Country	10 Clusters	25 Clusters
United States	1	1	**New Zealand**	9	20
China	2	2	**Algeria**	9	20
Japan	3	3	**Qatar**	9	20
Germany	4	4	**Hungary**	9	23
UK	5	5	**Kuwait**	9	23
France	6	6	**Angola**	9	23
India	7	7	**Morocco**	9	23
Italy	7	8	**Ecuador**	9	23
Brazil	7	8	**Sudan**	9	23
Canada	10	10	**Ukraine**	9	23

Korea	10	11	**Slovak**	9	23
Russia	10	11	**Sri Lanka**	9	23
Australia	10	11	**Oman**	9	24
Spain	10	14	**Dominican**	9	24
Mexico	10	15	**Uzbekistan**	9	24
Indonesia	8	16	**Guatemala**	9	24
Netherlands	8	17	**Kenya**	9	24
Turkey	8	17	**Myanmar**	9	24
Switzerland	8	19	**Ethiopia**	9	24
Saudi Arabia	8	19	**Luxembourg**	9	24
Argentina	8	21	**Belarus**	9	24
Sweden	8	22	**Costa Rica**	9	24
Nigeria	8	22	**Uruguay**	9	24
Poland	8	22	**Azerbaijan**	9	24
Belgium	8	22	**Panama**	9	24
Thailand	9	9	**Bulgaria**	9	24
Norway	9	9	**Croatia**	9	24
Austria	9	9	**Lebanon**	9	24
Arab Emirates	9	9	**Tanzania**	9	24
Egypt	9	12	**Tunisia**	9	24
South Africa	9	12	**Slovenia**	9	24
Hong Kong	9	12	**Lithuania**	9	24
Denmark	9	13	**Yemen, Rep.**	9	24

Malaysia	9	13	**Ghana**	9	24
Singapore	9	13	**Jordan**	9	24
Philippines	9	13	**Serbia**	9	24
Colombia	9	13	**Congo**	9	24
Ireland	9	13	**Uganda**	9	25
Pakistan	9	13	**Paraguay**	9	25
Chile	9	18	**Latvia**	9	25
Finland	9	18	**El Salvador**	9	25
Portugal	9	20	**Trinidad**	9	25
Bangladesh	9	20	**Estonia**	9	25
Greece	9	20	**Nepal**	9	25
Vietnam	9	20	**Zambia**	9	25
Peru	9	20	**Honduras**	9	25
Czech	9	20	**Cyprus**	9	25
Kazakhstan	9	20	**Afghanistan**	9	25
Iraq	9	20	**Cambodia**	9	25
Romania	9	20	**Iceland**	9	25

After classifying the countries into their clusters in the table above, there are some clusters that have more countries than the others. Therefore, to identify the potentials of the countries within their clusters, the ranking is based on the weighted factors of the developed model.

Table 18: The ranking of the countries within their own 25 clusters

Country	Ranking	25 Clusters	Country	Ranking	25 Clusters
United States	1	1	**New Zealand**	10	20
China	1	2	**Algeria**	11	20
Japan	1	3	**Qatar**	12	20
Germany	1	4	**Hungary**	1	23
UK	1	5	**Kuwait**	2	23
France	1	6	**Angola**	3	23
India	1	7	**Morocco**	4	23
Italy	1	8	**Ecuador**	5	23
Brazil	2	8	**Sudan**	6	23
Canada	1	10	**Ukraine**	7	23
Korea	1	11	**Slovak**	8	23
Russia	2	11	**Sri Lanka**	9	23
Australia	3	11	**Oman**	1	24
Spain	1	14	**Dominican**	2	24
Mexico	1	15	**Uzbekistan**	3	24
Indonesia	1	16	**Guatemala**	4	24
Netherlands	1	17	**Kenya**	5	24
Turkey	2	17	**Myanmar**	6	24
Switzerland	1	19	**Ethiopia**	7	24
Saudi Arabia	2	19	**Luxembourg**	8	24
Argentina	1	21	**Belarus**	9	24

Sweden	1	22	**Costa Rica**	10	24
Nigeria	2	22	**Uruguay**	11	24
Poland	3	22	**Azerbaijan**	12	24
Belgium	4	22	**Panama**	13	24
Thailand	1	9	**Bulgaria**	14	24
Norway	2	9	**Croatia**	15	24
Austria	3	9	**Lebanon**	16	24
Arab Emirates	4	9	**Tanzania**	17	24
Egypt	1	12	**Tunisia**	18	24
South Africa	2	12	**Slovenia**	19	24
Hong Kong	3	12	**Lithuania**	20	24
Denmark	1	13	**Yemen, Rep.**	21	24
Malaysia	2	13	**Ghana**	22	24
Singapore	3	13	**Jordan**	23	24
Philippines	4	13	**Serbia**	24	24
Colombia	5	13	**Congo**	25	24
Ireland	6	13	**Uganda**	1	25
Pakistan	7	13	**Paraguay**	2	25
Chile	1	18	**Latvia**	3	25
Finland	2	18	**El Salvador**	4	25
Portugal	1	20	**Trinidad**	5	25
Bangladesh	2	20	**Estonia**	6	25
Greece	3	20	**Nepal**	7	25

Vietnam	4	20	**Zambia**	8	25
Peru	5	20	**Honduras**	9	25
Czech	6	20	**Cyprus**	10	25
Kazakhstan	7	20	**Afghanistan**	11	25
Iraq	8	20	**Cambodia**	12	25
Romania	9	20	**Iceland**	13	25

5.6 Developed model sensitivity analysis and robustness

The developed model of the real-world sample can be tested by other similarity coefficient methods and different clustering algorithms to show the degree of flexibility of the developed model. The result clusters of each similarity coefficient method will be compared with the main method used in this research, which was the complete-linkage clustering with Euclidean distance and also will be compared with each other.

The clustering algorithms that will be implemented with the real-world sample are listed below in a sequence:

A. Method 1: Implementation of Euclidean distance with complete-linkage clustering
B. Method 2: Implementation of CityBlock with complete-linkage clustering
C. Method 3: Implementation of Euclidean distance with average-linkage clustering
D. Method 4: Implementation of CityBlock with average-linkage clustering

5.6.1 Method 1: Implementation of Euclidean distance with complete-linkage clustering

This was the main method of the research, discussed in detail in section 5.5.

5.6.2 Method 2: Implementation of CityBlock with complete-linkage clustering

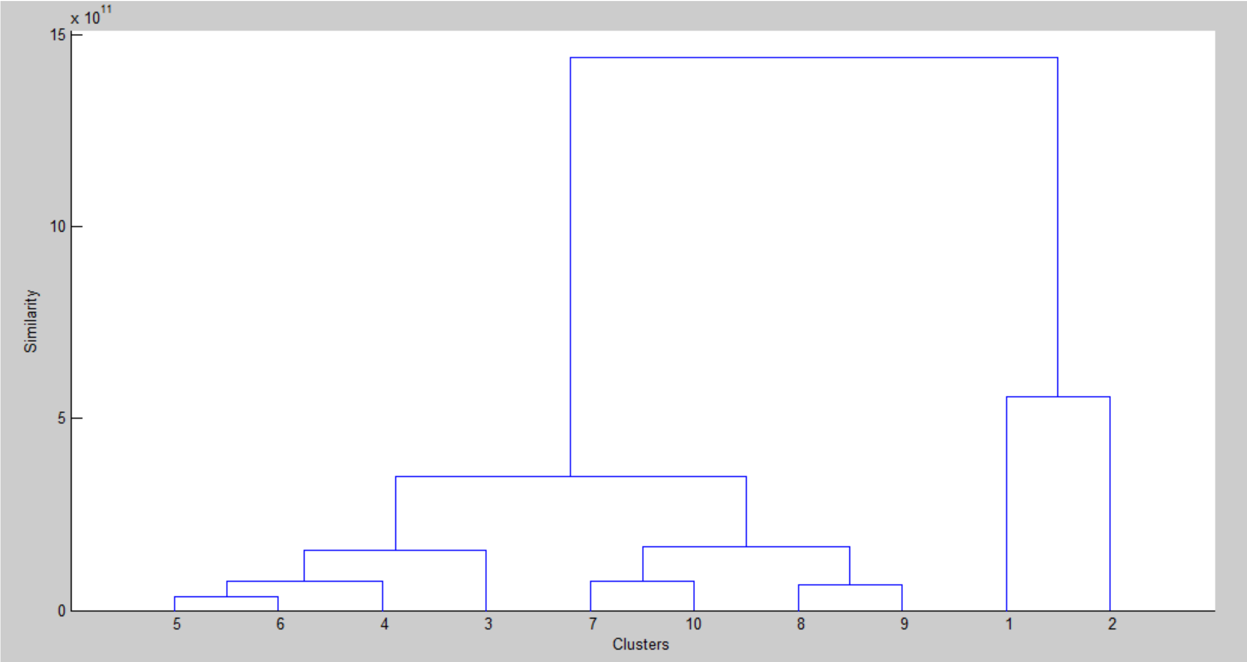

Figure 11: The distinct 10 clusters of the real-world sample of 100 countries by CityBlock with complete-linkage clustering

Table 19: The distinct clusters of the real-world sample of 100 countries by CityBlock with complete-linkage clustering for 10 and 25 clusters

Country	10Clusters	25Clusters	Country	10Clusters	25Clusters
United States	1	1	**New Zealand**	9	20
China	2	2	**Algeria**	9	20
Japan	3	3	**Qatar**	9	20
Germany	4	4	**Hungary**	9	23
United Kingdom	5	5	**Kuwait**	9	23
France	6	6	**Angola**	9	23
India	7	7	**Morocco**	9	23
Italy	7	8	**Ecuador**	9	23

Brazil	7	8	**Sudan**	9	23
Canada	10	10	**Ukraine**	9	23
Korea, Rep.	10	11	**Slovak Republic**	9	23
Russia	10	11	**Sri Lanka**	9	23
Australia	10	11	**Oman**	9	24
Spain	10	14	**Dominican**	9	24
Mexico	10	15	**Uzbekistan**	9	24
Indonesia	8	16	**Guatemala**	9	24
Netherlands	8	17	**Kenya**	9	24
Turkey	8	17	**Myanmar**	9	24
Switzerland	8	19	**Ethiopia**	9	24
Saudi Arabia	8	19	**Luxembourg**	9	24
Argentina	8	21	**Belarus**	9	24
Sweden	8	22	**Costa Rica**	9	24
Nigeria	8	22	**Uruguay**	9	24
Poland	8	22	**Azerbaijan**	9	24
Belgium	8	22	**Panama**	9	24
Thailand	9	9	**Bulgaria**	9	24
Norway	9	9	**Croatia**	9	24
Austria	9	9	**Lebanon**	9	24
Arab Emirates	9	9	**Tanzania**	9	24
Egypt	9	12	**Tunisia**	9	24

South Africa	9	12	**Slovenia**	9	24
Hong Kong	9	12	**Lithuania**	9	24
Denmark	9	13	**Yemen**	9	24
Malaysia	9	13	**Ghana**	9	24
Singapore	9	13	**Jordan**	9	24
Philippines	9	13	**Serbia**	9	24
Colombia	9	13	**Congo**	9	24
Ireland	9	13	**Uganda**	9	25
Pakistan	9	13	**Paraguay**	9	25
Chile	9	18	**Latvia**	9	25
Finland	9	18	**El Salvador**	9	25
Portugal	9	20	**Trinidad**	9	25
Bangladesh	9	20	**Estonia**	9	25
Greece	9	20	**Nepal**	9	25
Vietnam	9	20	**Zambia**	9	25
Peru	9	20	**Honduras**	9	25
Czech Republic	9	20	**Cyprus**	9	25
Kazakhstan	9	20	**Afghanistan**	9	25
Iraq	9	20	**Cambodia**	9	25
Romania	9	20	**Iceland**	9	25

5.6.3 Method 3: Implementation of Euclidean distance with average-linkage clustering

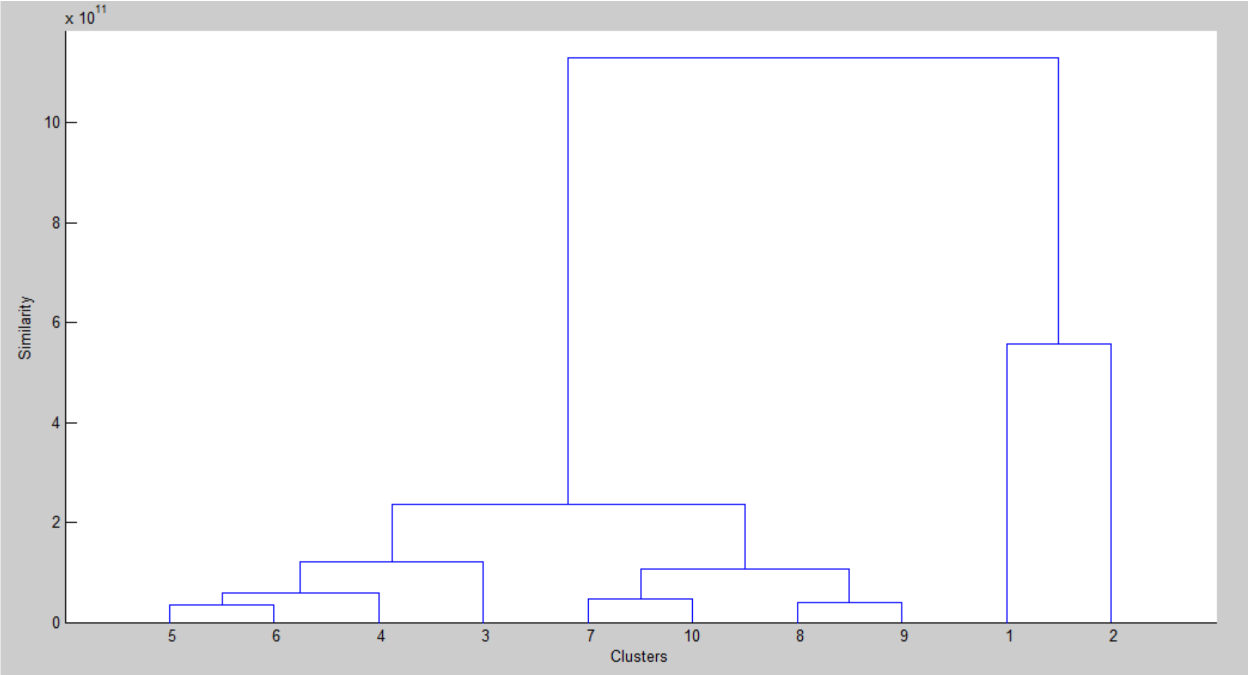

Figure 12: The distinct 10 clusters of the real-world sample of 100 countries by average-linkage clustering with Euclidean distance

Table 20: The distinct clusters of the real-world sample of 100 countries by average-linkage clustering with Euclidean distance for 10and 25 clusters

Country	10Clusters	25Clusters	Country	10Clusters	25Clusters
United States	1	1	**New Zealand**	9	23
China	2	2	**Algeria**	9	23
Japan	3	3	**Qatar**	9	23
Germany	4	4	**Hungary**	9	24
United Kingdom	5	5	**Kuwait**	9	24
France	6	6	**Angola**	9	24
India	7	7	**Morocco**	9	24
Italy	7	8	**Ecuador**	9	24
Brazil	7	8	**Sudan**	9	24

Canada	10	10	**Ukraine**	9	24
Korea, Rep.	10	11	**Slovak Republic**	9	24
Russia	10	11	**Sri Lanka**	9	24
Australia	10	13	**Oman**	9	25
Spain	10	14	**Dominican**	9	25
Mexico	10	15	**Uzbekistan**	9	25
Indonesia	8	16	**Guatemala**	9	25
Netherlands	8	17	**Kenya**	9	25
Turkey	8	17	**Myanmar**	9	25
Switzerland	8	19	**Ethiopia**	9	25
Saudi Arabia	8	19	**Luxembourg**	9	25
Argentina	8	21	**Belarus**	9	25
Sweden	8	22	**Costa Rica**	9	25
Nigeria	8	22	**Uruguay**	9	25
Poland	8	22	**Azerbaijan**	9	25
Belgium	8	22	**Panama**	9	25
Thailand	9	9	**Bulgaria**	9	25
Norway	9	9	**Croatia**	9	25
Austria	9	9	**Lebanon**	9	25
Arab Emirates	9	9	**Tanzania**	9	25
Egypt	9	12	**Tunisia**	9	25
South Africa	9	18	**Slovenia**	9	25
Hong Kong	9	18	**Lithuania**	9	25

Denmark	9	18	**Yemen, Rep.**	9	25
Malaysia	9	18	**Ghana**	9	25
Singapore	9	18	**Jordan**	9	25
Philippines	9	18	**Serbia**	9	25
Colombia	9	18	**Congo**	9	25
Ireland	9	18	**Uganda**	9	25
Pakistan	9	18	**Paraguay**	9	25
Chile	9	20	**Latvia**	9	25
Finland	9	20	**El Salvador**	9	25
Portugal	9	23	**Trinidad**	9	25
Bangladesh	9	23	**Estonia**	9	25
Greece	9	23	**Nepal**	9	25
Vietnam	9	23	**Zambia**	9	25
Peru	9	23	**Honduras**	9	25
Czech Republic	9	23	**Cyprus**	9	25
Kazakhstan	9	23	**Afghanistan**	9	25
Iraq	9	23	**Cambodia**	9	25
Romania	9	23	**Iceland**	9	25

5.6.4 Method 4: Implementation of CityBlock with average-linkage clustering

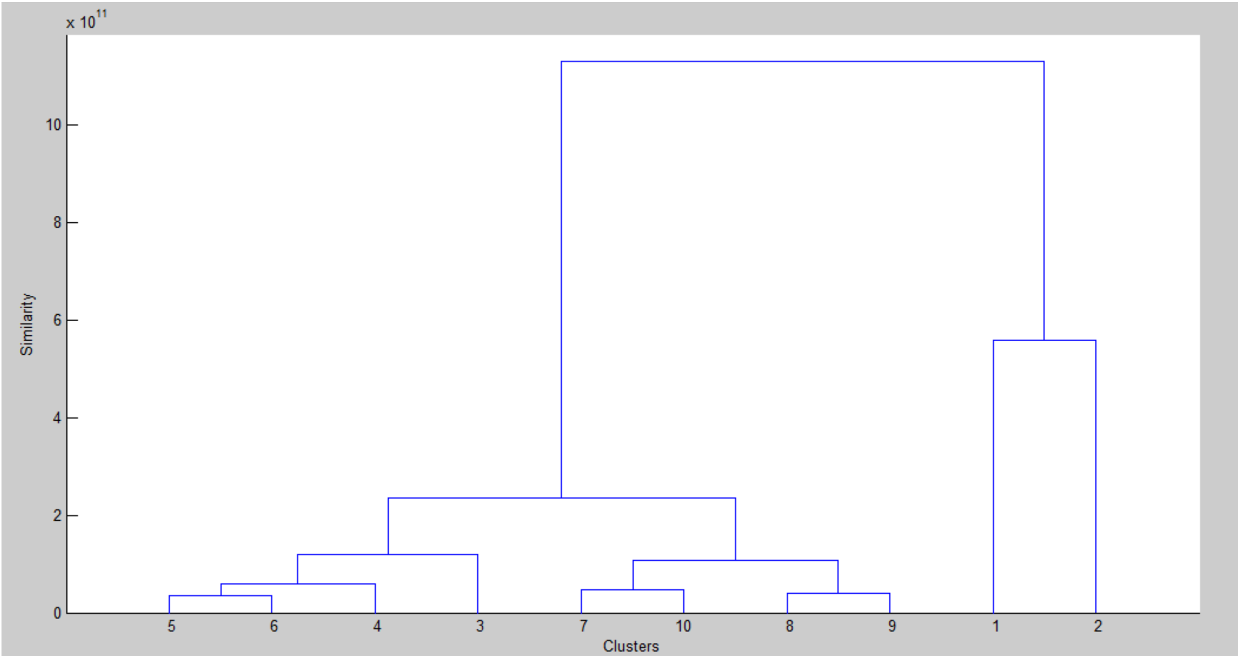

Figure 13: The distinct 10 clusters of the real-world sample of 100 countries by CityBlock with average-linkage clustering

Table 21: distinct clusters of the real-world sample of 100 countries by CityBlock with average-linkage clustering for 10 and 25 clusters

Country	10Clusters	25Clusters	Country	10Clusters	25Clusters
United States	1	1	**New Zealand**	9	23
China	2	2	**Algeria**	9	23
Japan	3	3	**Qatar**	9	23
Germany	4	4	**Hungary**	9	24
United Kingdom	5	5	**Kuwait**	9	24
France	6	6	**Angola**	9	24
India	7	7	**Morocco**	9	24
Italy	7	8	**Ecuador**	9	24
Brazil	7	8	**Sudan**	9	24

Canada	10	10	**Ukraine**	9	24
Korea, Rep.	10	11	**Slovak**	9	24
Russia	10	11	**Sri Lanka**	9	24
Australia	10	13	**Oman**	9	25
Spain	10	14	**Dominican**	9	25
Mexico	10	15	**Uzbekistan**	9	25
Indonesia	8	16	**Guatemala**	9	25
Netherlands	8	17	**Kenya**	9	25
Turkey	8	17	**Myanmar**	9	25
Switzerland	8	19	**Ethiopia**	9	25
Saudi Arabia	8	19	**Luxembourg**	9	25
Argentina	8	21	**Belarus**	9	25
Sweden	8	22	**Costa Rica**	9	25
Nigeria	8	22	**Uruguay**	9	25
Poland	8	22	**Azerbaijan**	9	25
Belgium	8	22	**Panama**	9	25
Thailand	9	9	**Bulgaria**	9	25
Norway	9	9	**Croatia**	9	25
Austria	9	9	**Lebanon**	9	25
Arab Emirates	9	9	**Tanzania**	9	25
Egypt	9	12	**Tunisia**	9	25
South Africa	9	18	**Slovenia**	9	25
Hong Kong	9	18	**Lithuania**	9	25

Denmark	9	18	**Yemen, Rep.**	9	25
Malaysia	9	18	**Ghana**	9	25
Singapore	9	18	**Jordan**	9	25
Philippines	9	18	**Serbia**	9	25
Colombia	9	18	**Congo**	9	25
Ireland	9	18	**Uganda**	9	25
Pakistan	9	18	**Paraguay**	9	25
Chile	9	20	**Latvia**	9	25
Finland	9	20	**El Salvador**	9	25
Portugal	9	23	**Trinidad**	9	25
Bangladesh	9	23	**Estonia**	9	25
Greece	9	23	**Nepal**	9	25
Vietnam	9	23	**Zambia**	9	25
Peru	9	23	**Honduras**	9	25
Czech Republic	9	23	**Cyprus**	9	25
Kazakhstan	9	23	**Afghanistan**	9	25
Iraq	9	23	**Cambodia**	9	25
Romania	9	23	**Iceland**	9	25

5.6.5 All the different clustering algorithms comparison

The method used in the research developed model was complete-linkage clustering (CLINK) with Euclidean distance that was selected for many reasons, as illustrated in Chapter Four. These are some of the advantages of using complete-linkage clustering with Euclidean distance:

- Prevents the merger of two clusters together for only a high level of similarity between two members while other members are dissimilar (chaining reaction problem of some clustering algorithms)
- Least similar pair between two clusters is used to determine the inter-cluster similarity
- Clusters are small and tightly bound
- Computer software-friendly (MATLAB)

Table 22: the four different methods of clustering algorithm implemented for 25 clusters

100 countries	Method 1	Method 2	Method 3	Method 4
United States	1	1	1	1
China	2	2	2	2
Japan	3	3	3	3
Germany	4	4	4	4
United Kingdom	5	5	5	5
France	6	6	6	6
India	7	7	7	7
Italy	8	8	8	8

Brazil	8	8	8	8
Canada	10	10	10	10
Korea, Rep.	11	11	11	11
Russian Federation	11	11	11	11
Australia	11	11	13	13
Spain	14	14	14	14
Mexico	15	15	15	15
Indonesia	16	16	16	16
Netherlands	17	17	17	17
Turkey	17	17	17	17
Switzerland	19	19	19	19
Saudi Arabia	19	19	19	19
Argentina	21	21	21	21
Sweden	22	22	22	22
Nigeria	22	22	22	22
Poland	22	22	22	22
Belgium	22	22	22	22
Thailand	9	9	9	9
Norway	9	9	9	9
Austria	9	9	9	9
United Arab Emirates	9	9	9	9
Egypt, Arab Rep.	12	12	12	12
South Africa	12	12	18	18

Hong Kong SAR, China	12	12	18	18
Denmark	13	13	18	18
Malaysia	13	13	18	18
Singapore	13	13	18	18
Philippines	13	13	18	18
Colombia	13	13	18	18
Ireland	13	13	18	18
Pakistan	13	13	18	18
Chile	18	18	20	20
Finland	18	18	20	20
Portugal	20	20	23	23
Bangladesh	20	20	23	23
Greece	20	20	23	23
Vietnam	20	20	23	23
Peru	20	20	23	23
Czech Republic	20	20	23	23
Kazakhstan	20	20	23	23
Iraq	20	20	23	23
Romania	20	20	23	23
New Zealand	20	20	23	23
Algeria	20	20	23	23
Qatar	20	20	23	23
Hungary	23	23	24	24

Kuwait	23	23	24	24
Angola	23	23	24	24
Morocco	23	23	24	24
Ecuador	23	23	24	24
Sudan	23	23	24	24
Ukraine	23	23	24	24
Slovak Republic	23	23	24	24
Sri Lanka	23	23	24	24
Oman	24	24	25	25
Dominican Republic	24	24	25	25
Uzbekistan	24	24	25	25
Guatemala	24	24	25	25
Kenya	24	24	25	25
Myanmar	24	24	25	25
Ethiopia	24	24	25	25
Luxembourg	24	24	25	25
Belarus	24	24	25	25
Costa Rica	24	24	25	25
Uruguay	24	24	25	25
Azerbaijan	24	24	25	25
Panama	24	24	25	25
Bulgaria	24	24	25	25
Croatia	24	24	25	25

Lebanon	24	24	25	25
Tanzania	24	24	25	25
Tunisia	24	24	25	25
Slovenia	24	24	25	25
Lithuania	24	24	25	25
Yemen, Rep.	24	24	25	25
Ghana	24	24	25	25
Jordan	24	24	25	25
Serbia	24	24	25	25
Congo, Dem. Rep.	24	24	25	25
Uganda	25	25	25	25
Paraguay	25	25	25	25
Latvia	25	25	25	25
El Salvador	25	25	25	25
Trinidad and Tobago	25	25	25	25
Estonia	25	25	25	25
Nepal	25	25	25	25
Zambia	25	25	25	25
Honduras	25	25	25	25
Cyprus	25	25	25	25
Afghanistan	25	25	25	25
Cambodia	25	25	25	25
Iceland	25	25	25	25

From the table above the four different methods of clustering algorithms implemented with the real-world sample yield a distinct cluster of each according to the attractiveness to a potential petrochemical facility location. The results clusters are listed, where the categories of the main clusters almost stay the same or with slight changes that demonstrate the robustness of the developed model of the research.

CHAPTER SIX

Conclusion

From the previous research on a similar goal of determining a new petrochemical facility location, the need of such an algorithm approach model arises because this approach features opportunities for the decision makers and the investors as well as exposing multiple leading factors in the petrochemical industry. The developed algorithm approach classified the countries in groups based on their similarities and dissimilarities of the attributes' data selected for the study of a potential petrochemical facility location. The countries were ranked based on their attractiveness level to the petrochemical manufacturing site in general as well as ranked within each cluster.

In this research, the petrochemical facility location problem has been solved by using similarity coefficient-based clustering algorithms that considered the decision factors taken from the previous research. This model approach suggests certain clusters of countries that are very similar in their attributes and behavior towards a petrochemical facility location. Moreover, in this developed model the leading factors to determine a petrochemical facility location have been quantified by real world numbers derived from the World Bank database website.

This developed approach reduces the error in data collection and the resulting analysis because it analyzed a huge amount of data for multiple factors. Decision makers who are identifying a new petrochemical facility location will have multiple options for similar countries to build a new plant despite other approaches that propose a single site or a ranking of locations. This is a very sensitive approach because any error will make a certain location lose the bid that might turn out to be based on inaccurate data or become a misleading guide.

For future research:

- More investigation might be needed to figure out the noise in the model, for example, the big influence of the data that has a huge impact on the clustering analysis.
- Use the model as one piece of a chain to create a trend for the petrochemical industry global market.

REFERENCES

Anderberg, M.R. 1973. *Cluster Analysis for Applications*.

Annett, A. 2001. “Social Fractionalization, Political Instability, and the Size of Government.” *IMF Staff Papers* 48 (3).

Ballance, Robert H. 1987. *International Industry and Business: Structural Change, Industrial Policy and Industry Strategies*. London: Allen & Unwin .

Balvers, Ronald, and Lázló Szerb. 1996. “Location in the Hotelling Duopoly Model with Demand Uncertainty.” *European Economic Review* 40 (7). North-Holland: 1453–61. doi:10.1016/0014-2921(95)00042-9.

Bass, Bernard M, Donald W Mcgregor, and James L Walters. 1977. “Selecting Foreign Plant Sites: Economic, Social and Political Considerations.” *Academy of Management Journal Dec* 20 (4): 1986–535.

Bater, J. & Walker, D. 1977. *Industrial Services: Literature and Research Prospects. In D. Walker (Ed.). Industrial Services*. Waterloo. University of Waterloo: Department of Geography.

Beckman, Sara Lynn, and Donald B. Rosenfield. 2008. *Operations Strategy: Competing in the 21st Century*. Irwin: McGraw-Hill.

Belli, R. David. 1971. “Sales of Foreign Affiliates of US Firms 1961-65, 1967 and 1968.” *Survey of Current Business* 50 (10): 18–20.

Bertil Ohlin. 1935. *Interregional and International Trade*.

Bertil Ohlin. 1952. *Interregional and International Trade*.

Button, Kenneth. 1996. *The Economics of Location Edited by Melvin L. Greenhut and George Norman*.

Carnoy, M. 1972. *Industrialization in a Latin American Common Market*. Washington DC: Brookings Institute.

Chernotsky, HI. 1983. "Selecting US Sites: A Case Study of German and Japanese Firms." *Management International Review*. http://www.jstor.org/stable/40227676.

Chisholm, M. 1971. "Freight Transport Costs, Industrial Location and Regional Development. In M. Chisholm and G. Manners (Eds). Spatial Policy Problems of the British Economy." *Cambridge: Cambridge University Press*, 213–44.

Coughlin, C.; Joseph, V. & Vachira, A. 1990. "State Government Effects on the Location of Foreign Direct Investment." *Regional Science Perspectives* 20: 194–207.

Dicken, P. & Lloyd, P. 1978. "Inner Metropolitan Industrial Change, Enterprise Structure and Policy Issues: Case Studies of Manchester and Merseyside." *Regional Studies* 12: 181–97.

Dicken, P. 1986. *Global Shift: Industrial Change in a Turbulent World*. London: Harper and Row.

Dicken, Peter. 2011. *Global Shift: Mapping the Changing Contours of the World Economy*. Guilford Press.

Dziembowska-Kowalska, J. & Funck, R. 2000. "Cultural Activities as a Location Factor in

European Competition between Regions: Concepts and Some Evidence." *The Annals of Regional Science* 34 (1): 1–12.

Epping, G. Micheal. 1982. "Important Factors in Plant Location in 1980" Vol. 13 (Issue 2).

Flaherty, M. 1996. *Global Operations Management*.

Friedman, A. 1977. *Industry and Labor: Class Struggle at Work and Monopoly Capitalism.* London: Macmillan.

Galan, J. & Gonzalez-Benito, J. 2006. "Distinctive Determinant Factors of Spanish Foreign Investment in Latin America." *Journal of World Business* 41: 171–89.

Gold, S. 1991. "A New Approach to Site Selection." *Distribution* 90: 29–33.

Greenhut, M. 1956. *Plant Location in Theory and Practice. Chapel Hill, NC: University of North Carolina Press.*

Haitani, K. & Marquis, C. 1990. "Japanese Investment in the Southeast United States: Factors, Obstacles, and Opportunities." *Economic Development Review* 8: 42–49.

Hanson, Gordon H. 2012. "The Rise of Middle Kingdoms: Emerging Economies in Global Trade." *The Journal of Economic Perspectives* 26 (2): 41–63.

Ho, W. 2008. "Integrated Analytic Hierarchy Process and Its applications–A Literature Review." *European Journal of Operational Research*. https://pdfs.semanticscholar.org/5eb4/0dbde1acfb7286fafa1f3923be7ef4162d2a.pdf.

Hoover, Edgar. 1937. *Location Theory and the Shoe Leather Industries*.

Hoover, Edgar. 1948. *The Location of Economic Activity*. http://krishikosh.egranth.ac.in/handle/1/2025674.

Horst, Thomas. 1972. "Firm and Industry Determinants of the Decision to Invest Abroad: An Empirical Study AN EMPIRICAL STUDY*." *Source: The Review of Economics and Statistics* 54 (3). The MIT Press: 258–66. http://www.jstor.org/stable/1937986.

Hotelling, Harold. 1929. "Stability in Competition." *The Economic Journal* 39 (153): 41–57.

Isard, Walter. 1956. *Location and Space-Economy*.

Jaccard, Paul. 1908. "Nouvelles Recherches Sur La Distribution Florale." *Bulletin de La Société Vaudoise de Sciences Naturelles* 44: 223–370.

Jungthirapanich, Chamnong, and Colin O. Benjamin. 1995. "A Knowledge-Based Decision Support System for Locating a Manufacturing Facility.Pdf." *IIE Transactions* 27 (6): 789–99.

Katz, A. 1995. "More on Hotelling's Stability in Competition." *International Journal of Industrial Organization* 13: 89–93.

Kinkel, Steffen. 2012. "Trends in Production Relocation and Backshoring Activities: Changing Patterns in the Course of the Global Economic Crisis." *International Journal of Operations & Production Management* 32 (6): 696–720.

Launhardt, W. 1885. *Mathematische Begrundung Der Volkwirthschaftslehre*.

Lerner, A P, and H W Singer. 1937. "Some Notes on Duopoly and Spatial Competition." *Source Journal of Political Economy* 45 (2): 145–86. http://www.jstor.org/stable/1824516.

Losch, A. 1954. “The Economics of Location. New Haven: Yale University Press.”

Lösch, August. 1939. *Eine Neue Theorie Des Internationalen Handels*.

Lowe, J. & Moryadas, S. 1975. *The Geography of Movement. Boston: Houghton Mifflin.*

Maccarthy, B L, and W Atthirawong. 2003. “Factors Affecting Location Decisions in International Operations --a Delphi Study.” *International Journal of Operations & Production Management* 23 (78).

Malecki, E. 1984. “High Technology and Local Economic Development.” *APA Journal* 8: 262–69.

Mcauley, John. 1972. “Machine Grouping for Efficient Production T.”

Mckinnon, A. 1989. *Physical Distribution Systems.* London: Routledge.

McKinnon, A. 1983. “The Development of Warehousing in Engkand.” *Geoforum* 14: 389–99.

McMillan, T. Jr. 1965. “Why Manufacturers Choose Plant Location vs Determinants of Plant Location.” *Land Economics* 43: 239–46.

Moriarty, B. 1980. *Industrial Location and Community Development.* Chapel Hill, N.C: University of North Carolina Press.

Noyelle, T. & Stanback, T. 1984. *The Economic Transformation of American Cities.* Totowa, NJ: Rowman and Allanheld.

Pietlock, B. 1992. “Developing Foreign Location Factors.” *Cost Engineering* 34: 7–11.

Piper, James R. 1971. “How US Firms Evaluate Foreign Investment Opportunities.” *MSU*

BUSINESS TOPICS-MICHIGAN STATE UNIVERSITY 19 (3): 11–20.

Rees, J. 1972. “The Industrial Corporation and Location Decision Analysis.” *Area* 4: 199–204.

Rezazadeh, Mehrdad, and Sajjad Farahani. 2010. “A Multi-Objective Approach for Multi Capacity Warehouse Location within Distribution Supply Chain Problem.” *Journal of American Science*.

Rummel, R], and David A Heenan. 1978. “How Multinationals Analyze Political Risk.” *Harvard Business* 54 (1): 90–105.

Sant, M. 1975. *Industrial Movement and Regional Development: The British Case.* Oxford: Pergamon.

Saxenian, A. 1985. “The Genesis of Silicon Valley. In P. Hall and A. Markusen (Eds). Silicon Landscapes.” *Allen and Unwin*, 20–34.

Schmenner, R. 1982. *Making Business Location Decisions*. Englewood Cliffs, NJ: Prentice Hall.

Seifoddini, Hamid, a Manocher Djassemi. 1995. “Merits of the Production Volume Based Sinlilarity Coefficient in Machine Cell Formation.” *Industrial Technology* 26.

Seifoddini, Hamid, and Manucher Djassemi. 1991. “The Production Data-Based Similarity Coefficient versus Jaccard’s Similarity Coefficient.” *Computers & Industrial Engineering 21* 1: 263–66.

Seifoddini, Hamid K. 1989. “Single Linkage versus Average Linkage Clustering in Machine Cells Formation Applications.” *Computers & Industrial Engineering*, 419–26.

Siebert, Horst. 2006. “Locational Competition: A Neglected Paradigm in the International

Division of Labour." *World Economy*. doi:10.1111/j.1467-9701.2006.00775.x.

Smith-Hamilton, A. & Omar, M. 2005. "FDI, International Business and Regulation - The Behaviour of UK Multinational Corporations." *European Business* 17 (1): 96–82.

Smithies, A. 1941. "OPTIMUM LOCATION IN SPATIAL COMPETITION." *Journal of Political Economy* 69.

Sneath, PHA, and RR Sokal. 1973. *Numerical Taxonomy. The Principles and Practice of Numerical Classification.* http://www.cabdirect.org/abstracts/19730310919.html.

Sokal, RR, and PHA Sneath. 1961. "Principles of Numerical Taxonomy." http://philpapers.org/rec/SOKPON.

Tellier, Normand , and Claude Vertefeuille. 1995. "UNDERSTANDING SPATIAL INERTIA: CENTER OF GRAVITY, POPULATION DENSITIES, THE WEBER PROBLEM, AND GRAVITY POTENTIAL." *Journal of Regional Science*. doi:10.1111/j.1467-9787.1995.tb01404.x.

Thisse, J. & Wildasin, D. 1995. "Optimal Transportation Policy with Strategic Locational Choice." *Regional Science & Urban Economics* 25: 395–410.

Thunen, J. 1875. *Der Isolirte Staat in Beziehung Auf Landwirthschaft Und Nationalökonomie*.

Tomback, M. (1995). 1995. "Multinational Plant Location as a Game of Timing." *European Journal of Operational Research*, no. 86: 434–51.

Townroe, P. 1969. "Locational Choice and the Individual Firm." *Regional Studies* 3: 15–24.

Vastag, Gyula, Sándor Kerekes, and Dennis A. Rondinelli. 1996. "Evaluation of Corporate

Environmental Management Approaches: A Framework and Application.” *International Journal of Production Economics* 43 (2–3): 193–211.

Venables, A. 1996. “Equilibrium Locations of Vertically Linked Industries.” *International Economic Review* 37: 341–59.

Vernon, R. 1968. *Manager in the International Economy*. *Englewood Cliffs: Prentice Hall*.

Vernon, R. 1971. *Sovereignty at Bay: The Multinational Spread of US Enterprises*. New York: Basic Books.

Walters, B. & Wheeler, J. 1984. “Localization Economies in the American Carpet Industry.” *Geographical Review* 74: 183–91.

Weber, Alfred. 1929. *Theory of the Location of Industries*.

Wheeler, D. & Mody, A. 1992. “International Investment Location Decisions: The Case of U.S. Firms.” *Journal of International Economics* 33: 57–76.

Young, G. 1994. “International Competitiveness, International Taxation and Domestic Investment.” *National Institute of Economic Review* 148: 44–48.

APPENDIX

The model data of 28 attributes and 100 countries

Country Name	Industry, value added	Population	Services	Lending interest rate	Cost of business start-up procedures	Wage and salaried workers	Lead time to import	Internet Users
Ranking Weight	0.04	0.02	0.04	0.01	0.04	0.05	0.01	0.02
United States	1.706	321418820	77.980	3.260	1.1	93.540	3	74.452
China	6.219	1371220000	50.236	4.350	0.7	86.300	5	50.300
Japan	1.492	126958472	73.415	1.143	7.5	88.460	3	91.058
Germany	1.333	81679769	68.877	15.407	2	89.250	3	87.590
UK	2.194	65128861	79.941	0.500	0.1	84.760	3	92.000
France	0.918	66538391	78.760	6.600	0.8	88.410	1	84.695
India	8.220	1311050527	52.969	10.008	14.3	18.100	5	26.000
Italy	0.812	60730582	74.224	4.129	14	75.620	3	65.572
Brazil	-6.326	207847528	72.683	43.958	4.8	67.520	4	59.079
Canada	-2.109	35848610	69.312	2.775	0.4	84.620	2	88.470
Korea, Rep.	1.715	50617045	59.707	3.533	14.5	73.210	3	89.649
Russia	-2.436	144096870	62.653	15.717	1.1	92.790	7	70.099
Australia	1.625	23789752	71.946	5.575	0.7	82.940	2	84.561
Spain	3.979	46443994	73.800	15.407	5.2	82.690	4	78.690
Mexico	0.949	127017224	63.615	3.423	18.1	67.890	3	57.431
Indonesia	2.653	257563815	46.464	12.663	19.9	38.700	5	21.976
Netherlands	0.481	16939923	78.168	1.650	4.6	83.160	2	93.097
Turkey	3.331	78665830	64.959	15.407	16.6	67.140	2	53.745
Switzerland	-0.676	8281430	73.800	2.681	2.3	85.780	2	87.479
Saudi Arabia	3.988	31540372	51.838	15.407	4.1	95.180	7	69.616
Argentina	1.469	43416755	65.891	24.916	9.7	74.970	4	69.401
Sweden	5.217	9799186	72.399	3.314	0.5	89.740	1	90.610
Nigeria	-2.243	182201962	58.760	16.849	31.7	88.300	3	47.443
Poland	5.795	37986412	63.255	5.500	12.2	78.770	1	67.997
Belgium	3.800	11249420	77.074	9.500	5	84.820	3	85.053
Thailand	2.219	67959359	55.140	6.563	6.7	45.610	1	39.316

Norway	1.445	5190239	63.539	4.280	0.9	92.950	2	96.810
Austria	0.949	8638366	70.369	5.600	0.3	87.000	2	83.926
Arab Emirates	4.384	9156963	61.719	15.407	11.2	95.820	2	91.243
Egypt	1.031	91508084	52.499	11.625	7.6	62.520	3	37.819
South Africa	1.126	55011976	68.730	9.417	0.3	85.420	3	51.919
Hong Kong	0.990	7305700	92.447	5.000	1.2	91.040	3	84.948
Denmark	1.908	5683483	75.844	7.100	0.2	91.350	1	96.331
Malaysia	5.244	30331007	55.117	4.585	6.7	73.900	7	71.064
Singapore	-3.404	5535002	73.561	5.350	0.6	86.010	2	82.103
Philippines	6.048	100699395	58.963	5.578	16.1	59.260	7	40.700
Colombia	1.962	48228704	59.188	11.450	7.5	49.040	3	55.905
Ireland	90.424	4643740	57.333	2.650	0.2	82.830	2	80.122
Pakistan	4.812	188924874	54.934	15.407	13.9	35.970	5	18.000
Chile	1.201	17948141	63.312	5.515	0.7	70.170	1	64.289
Finland	-1.106	5479531	70.610	3.400	1	85.780	2	92.651
Portugal	0.804	10358076	75.416	5.200	2.2	81.580	8	68.633
Bangladesh	9.667	160995642	56.346	11.709	13.9	29.410	5	14.400
Greece	0.689	10820883	80.192	6.800	2.2	65.040	3	66.835
Vietnam	9.639	91713300	44.160	7.118	4.9	39.310	3	52.720
Peru	1.588	31376670	59.408	16.106	9.8	47.070	1	40.900
Czech	4.306	10546059	59.705	4.283	6.7	82.660	5	81.299
Kazakhstan	-0.224	17544126	62.509	15.407	0.5	72.990	3	70.830
Iraq	9.510	36423395	61.719	12.293	37.4	67.421	7	17.220
Romania	6.560	19815308	60.340	6.768	2.1	71.020	3	55.763
New Zealand	4.524	4595700	70.891	5.763	0.3	85.080	2	88.223
Algeria	3.168	39666519	48.443	8.000	10.9	58.200	5	38.200
Qatar	1.747	2235355	41.338	4.436	5.1	99.580	3	92.885
Hungary	7.145	9843028	63.968	2.902	7.5	88.990	2	72.835
Kuwait	-1.668	3892115	48.304	4.295	2.3	96.770	1	82.079
Angola	3.796	25021974	26.960	16.882	22.5	67.421	14	12.400
Morocco	2.753	34377511	56.299	11.500	9.1	45.000	5	57.080
Ecuador	-0.871	16144363	55.808	9.800	22	56.190	3	48.940
Sudan	4.716	40234882	58.080	19.200	14.8	67.421	12	26.615
Ukraine	-13.393	45154029	59.661	21.823	0.6	84.070	2	48.885
Slovak	7.669	5423801	61.517	5.800	1.5	84.840	2	77.635
Sri Lanka	2.965	20966000	60.610	7.400	18.7	56.440	2	29.989
Oman	4.852	4490541	46.047	4.762	3.2	96.280	2	74.174
Dominican	8.517	10528391	66.910	14.877	16.4	55.550	4	54.216
Uzbekistan	8.453	31298900	47.110	15.407	3.4	67.421	20	42.800
Guatemala	3.737	16342897	60.809	13.228	25.1	44.140	3	27.100
Kenya	6.872	46050302	47.537	16.087	35.5	95.840	3	45.623

Myanmar	8.722	53897154	38.707	13.000	97.1	67.421	1	21.800
Ethiopia	21.656	99390750	42.774	8.000	79.1	48.260	14	11.600
Luxembourg	-2.319	569604	87.677	5.300	2	89.940	2	97.334
Belarus	-6.798	9489616	52.230	18.700	0.9	94.200	4	62.230
Costa Rica	3.415	4807850	72.325	14.233	11.1	75.690	4	59.763
Uruguay	1.062	3431555	64.057	15.841	21.7	73.020	3	64.600
Azerbaijan	1.600	9649341	56.216	17.534	1.2	32.160	7	77.000
Panama	5.975	3929141	69.367	7.456	6.3	67.310	4	51.205
Bulgaria	4.159	7177991	67.341	7.461	1.4	87.970	2	56.656
Croatia	1.883	4203604	69.504	9.200	3.4	84.450	2	69.803
Lebanon	0.739	5850743	78.597	7.091	34.2	62.090	1	74.000
Tanzania	11.260	53470420	42.873	16.105	23.2	13.400	4	5.355
Tunisia	-1.472	11253554	61.353	4.800	3.9	71.970	3	48.520
Slovenia	1.013	2063531	64.870	5.900	0	83.500	2	73.099
Lithuania	1.145	2904910	66.521	5.900	0.6	87.730	3	71.378
Yemen, Rep.	3.796	26832215	61.719	22.000	68	41.620	7	25.100
Ghana	0.997	27409893	51.415	25.000	19.4	22.500	4	23.478
Jordan	2.226	7594547	66.185	8.477	20.7	84.560	7	53.400
Serbia	3.046	7095383	60.455	14.800	6.6	69.840	2	65.317
Congo	4.801	77266814	46.910	19.371	29.3	67.421	7	3.800
Uganda	7.754	39032383	52.799	22.601	39.7	19.610	6	19.221
Paraguay	2.356	6639123	51.547	19.736	39.9	55.790	1	48.439
Latvia	2.212	1977527	73.722	5.900	1.5	87.390	1	79.201
El Salvador	3.160	6126583	61.853	8.100	42.7	53.450	2	26.915
Trinidad	-2.997	1360088	58.890	8.183	0.7	76.900	7	69.198
Estonia	-0.147	1314608	69.172	4.481	1.3	90.650	1	88.407
Nepal	1.478	28513700	51.555	8.000	28.4	24.630	3	17.582
Zambia	6.811	16211767	59.435	13.250	34.3	20.770	6	21.000
Honduras	3.181	8075060	58.778	20.658	42	46.490	7	20.357
Cyprus	0.888	1165300	87.169	6.700	12.5	85.070	1	71.716
Afghanistan	4.107	32526562	55.001	15.000	19	67.421	5	8.260
Cambodia	11.728	15577899	42.333	15.407	60.7	44.400	4	19.000
Iceland	4.411	330815	70.438	7.610	2.2	87.660	1	98.200

Country Name	Imports of goods and services	Industry, value added	Exports of goods and services	Time required to get electricity (days)	Time required to start a business	Labor force	Unemployment %
Ranking Weight	0.04	0.04	0.04	0.04	0.04	0.05	0.04
United States	15.448	20.685	12.554	89.6	5.6	162846081	4.906
China	18.489	40.932	21.973	143.2	28.9	806460138	4.605
Japan	17.959	25.473	17.636	97.7	11.2	65233384	3.136
Germany	39.213	30.488	46.781	28	10.5	42867313	4.311
United Kingdom	29.221	19.407	27.631	79	4.5	33747423	4.849
France	31.405	19.504	30.028	71	3.5	29921021	9.965
India	22.266	29.581	19.953	45.9	26	511066647	3.458
Italy	26.981	23.528	30.058	124	6.5	25322859	11.541
Brazil	14.067	22.345	12.889	64.4	79.5	108563800	11.452
Canada	33.969	28.845	31.580	137	1.5	19883470	7.073
Korea, Rep.	38.942	37.983	45.901	18	4	26661449	3.652
Russian Federation	20.604	32.789	28.667	160.5	9.8	75568612	5.723
Australia	21.209	25.552	19.797	75	2.5	12654298	5.738
Spain	30.728	23.638	33.178	107	13	23043451	19.447
Mexico	37.473	32.778	35.356	100.4	8.4	58284631	4.015
Indonesia	20.846	40.015	21.092	57.7	24.9	127149618	5.600
Netherlands	71.684	20.028	82.463	110	4	9034299	6.166
Turkey	30.841	26.512	27.961	63	6.5	29875461	10.329
Switzerland	51.189	25.519	62.903	39	10	4876026	4.583
Saudi Arabia	38.778	45.898	33.750	61	16.2	12566798	5.529
Argentina	11.848	28.065	11.008	92	25	20052816	6.557
Sweden	40.843	26.283	45.590	52	7	5248808	7.093
Nigeria	10.790	20.382	10.657	195.2	25.2	59126465	5.005
Poland	46.455	34.143	49.551	122	37	18322217	6.183
Belgium	81.258	22.179	82.928	88	4	5027011	8.256
Thailand	57.742	35.717	69.059	37	25.5	40069866	0.626
Norway	31.960	34.648	37.386	66	4	2789255	4.806
Austria	49.052	28.341	53.060	23	21	4473700	6.109
United Arab Emirates	83.116	28.686	97.359	28	8.2	6328674	3.691
Egypt, Arab Rep.	21.648	36.320	13.207	54	6.5	30786696	12.014
South Africa	31.730	29.441	30.723	84	43	21119236	25.927
Hong Kong	199.268	7.482	201.606	27	1.5	3870961	3.422
Denmark	47.826	22.926	55.227	38	3	2948290	6.051

Malaysia	63.251	36.430	70.904	31	18.5	14398343	3.298
Singapore	149.622	26.401	176.495	30	2.5	3189210	1.829
Philippines	34.846	30.771	28.189	42	28	45209579	5.876
Colombia	24.236	33.985	14.715	109	9	25402841	9.872
Ireland	92.248	41.654	123.995	85	5	2186580	8.089
Pakistan	17.019	19.959	10.585	180.7	18	67918996	5.870
Chile	30.253	32.810	29.980	43	5.5	9071864	6.566
Finland	37.139	26.932	36.841	42	14	2677385	8.997
Portugal	39.826	22.251	40.555	41	4.5	5187937	11.160
Bangladesh	24.749	28.146	17.337	428.9	19.5	72027272	4.066
Greece	31.771	15.691	31.915	51	13	4770478	23.909
Vietnam	88.988	36.955	89.779	46	24	55919519	2.178
Peru	23.710	32.835	21.298	67	26	17068996	4.946
Czech Republic	76.825	37.773	82.955	68	9	5323282	4.045
Kazakhstan	24.685	32.526	28.456	77	9	9222121	5.228
Iraq	22.077	28.686	34.836	56	34.5	9448523	16.045
Romania	41.619	34.901	41.094	182	12	9278282	6.421
New Zealand	27.264	23.009	27.264	58	0.5	2497244	5.246
Algeria	36.789	38.888	23.564	180	20	12584544	11.222
Qatar	36.000	58.502	56.056	90	8.7	1625720	0.227
Hungary	81.816	31.900	90.727	257	7	4527564	5.168
Kuwait	45.263	51.065	54.379	64	61.4	2133472	2.439
Angola	36.143	64.884	33.927	145	36	9257377	6.579
Morocco	42.085	29.224	34.294	49	9.5	12510890	9.983
Ecuador	23.661	34.078	20.674	74	48.5	7506619	5.361
Sudan	10.919	2.595	8.182	70	36.5	11861785	13.302
Ukraine	54.763	26.302	52.769	281	5	22554261	8.874
Slovak Republic	91.063	34.820	93.488	121	11.5	2736798	9.991
Sri Lanka	27.952	30.710	20.526	100	9	8240522	4.981
Oman	52.507	52.380	56.086	62	6.3	2541679	17.522
Dominican Republic	29.283	27.303	24.720	67	14.5	4893978	14.363
Uzbekistan	22.174	34.626	20.665	89	5.5	14046754	8.895
Guatemala	30.043	28.059	21.290	39	19.5	6565731	2.377
Kenya	29.038	19.524	15.769	97	22	18533603	10.998
Myanmar	26.537	34.545	20.780	77	13	30889201	0.806
Ethiopia	27.346	16.253	9.827	95	35	49975441	5.735
Luxembourg	202.571	12.079	235.585	56	16.5	284134	5.942
Belarus	59.828	39.974	60.022	105	5	4787557	0.530
Costa Rica	32.246	22.187	30.545	45	22.5	2335370	8.999
Uruguay	22.597	28.895	22.302	48	6.5	1774194	8.154
Azerbaijan	34.825	36.993	37.814	69	3	4928392	5.073

Panama	61.413	27.741	61.413	35	6	1907718	5.815
Bulgaria	63.959	27.872	64.107	130	23	3309019	8.001
Croatia	46.639	26.205	49.378	65	7	1862041	13.476
Lebanon	64.967	16.587	56.964	75	15	2158848	6.784
Tanzania	26.344	26.063	21.621	109	26	23763063	2.619
Tunisia	51.781	28.199	40.792	65	11	4149938	14.786
Slovenia	68.819	32.746	77.936	38	7	1005029	8.693
Lithuania	76.513	29.844	75.863	85	5.5	1460385	9.185
Yemen, Rep.	22.486	50.351	10.241	80	40.5	8223642	17.057
Ghana	55.398	27.597	43.850	79	14	13228560	5.766
Jordan	60.490	29.641	37.588	50	12.5	2008909	13.241
Serbia	56.426	31.361	46.674	125	7	3045429	16.530
Congo, Dem. Rep.	34.973	32.464	29.490	54	11.5	30692140	3.636
Uganda	28.765	21.381	17.710	66	26	17872336	2.281
Paraguay	41.778	29.794	42.677	67	35	3383114	5.437
Latvia	60.112	23.116	58.980	107	5.5	1010941	9.883
El Salvador	42.030	26.868	25.959	59	15.5	2844297	6.287
Trinidad and Tobago	43.104	40.578	41.114	61	10.5	678263	3.853
Estonia	75.137	27.443	79.286	91	3.5	680578	6.911
Nepal	41.663	15.442	11.675	70	17	16333723	3.213
Zambia	47.176	35.314	37.140	117	8.5	6834809	7.530
Honduras	62.589	27.534	44.846	39	13	3715178	6.284
Cyprus	60.883	10.556	61.218	137	6	629995	11.728
Afghanistan	49.830	23.283	7.150	114	7.5	9936806	8.540
Cambodia	66.146	29.417	61.718	179	99	8789877	0.265
Iceland	46.176	23.321	53.703	22	3.5	196582	3.760

Country Name	Quality of port infrastructure	Air transport, freight (million ton-km)	Manufacturing, value added (% of GDP)	Trade (% of GDP)	Foreign direct investment	Trade in services (% of GDP)
Ranking Weight	0.04	0.01	0.04	0.03	0.06	0.04
United States	5.7	37218.889	12.376	28.002	2.104	6.872
China	4.5	19805.630	29.738	40.462	2.258	6.827
Japan	5.4	8868.745	17.707	35.595	-0.001	7.733
Germany	5.6	6985.008	22.809	85.994	1.374	16.721
United Kingdom	5.7	5466.505	9.761	56.852	2.043	19.401
France	5.3	4098.310	11.231	61.432	1.446	19.533
India	4.2	1833.848	16.585	42.219	2.107	14.497
Italy	4.3	945.434	15.788	57.039	0.714	10.900
Brazil	2.7	1493.939	11.763	26.956	4.162	5.792
Canada	5.5	2074.831	10.623	65.549	3.523	11.273
Korea	5.2	11296.967	29.489	84.843	0.366	15.347
Russia	3.9	4761.047	13.769	49.271	0.474	10.276
Australia	5	1887.296	6.832	41.007	2.885	7.975
Spain	5.7	1040.913	14.236	63.906	2.121	15.348
Mexico	4.3	713.985	18.444	72.829	2.873	4.804
Indonesia	3.8	747.473	20.843	41.938	2.327	6.409
Netherlands	6.8	5292.795	11.700	154.147	13.567	39.032
Turkey	4.5	2882.162	17.644	58.802	2.377	9.647
Switzerland	4.6	1322.379	17.960	114.092	14.547	30.641
Saudi Arabia	4.8	1783.086	12.300	72.528	1.260	16.203
Argentina	3.8	243.773	17.252	22.857	2.049	5.436
Sweden	5.6	256.516	17.001	86.434	1.613	26.605
Nigeria	3	22.401	9.532	21.447	0.643	4.776
Poland	4	120.016	19.694	96.006	2.949	16.287
Belgium	6.3	1464.317	14.269	164.186	-4.570	48.046
Thailand	4.5	2134.149	26.917	126.801	2.278	28.520
Norway	5.5	176.585	8.034	69.346	-2.255	22.434
Austria	4	351.379	18.906	102.112	1.141	27.524
Arab Emirates	6.5	16647.479	14.234	180.475	2.375	25.328
Egypt, Arab Rep.	4.3	397.532	16.587	34.855	2.081	13.070
South Africa	4.9	885.278	13.218	62.452	0.484	9.723
Hong Kong	6.4	11294.272	1.189	400.874	58.547	57.695
Denmark	5.8	1.252	14.632	103.053	0.627	38.355
Malaysia	5.6	2005.979	22.791	134.155	3.700	25.206

Singapore	6.7	6154.365	19.810	326.117	22.294	96.700
Philippines	3.2	484.191	20.058	63.035	1.995	18.207
Colombia	3.6	1317.562	12.209	38.951	4.017	6.316
Ireland	5.3	138.580	36.927	216.243	71.717	106.49
Pakistan	4.1	183.177	13.416	27.604	0.361	5.172
Chile	4.9	1392.236	11.927	60.233	8.496	9.704
Finland	6.4	713.484	16.962	73.980	7.326	21.830
Portugal	5.3	343.971	13.763	80.381	0.318	21.238
Bangladesh	3.6	182.693	17.611	42.086	1.733	5.544
Greece	4.6	27.453	9.483	63.686	0.586	22.157
Vietnam	3.9	384.470	15.220	178.767	6.095	13.791
Peru	3.6	223.643	14.501	45.008	4.134	7.310
Czech Republic	3.6	26.620	26.973	159.780	1.339	22.888
Kazakhstan	2.9	37.669	10.832	53.141	3.571	8.762
Iraq	2.7	10.758	14.234	56.913	1.842	7.396
Romania	3.4	4.691	23.651	82.712	2.426	16.510
New Zealand	5.5	999.385	11.864	91.719	-0.078	15.053
Algeria	3	24.723	14.234	60.354	-0.245	8.819
Qatar	5.6	7563.307	9.688	92.056	0.650	27.801
Hungary	3.4	0.000	24.592	172.543	-2.156	31.359
Kuwait	4	275.778	6.155	99.642	0.250	26.177
Angola	2.7	46.043	14.234	70.070	9.045	20.989
Morocco	4.8	47.828	18.026	76.379	3.234	22.454
Ecuador	4.8	86.129	15.704	44.334	1.058	5.588
Sudan	2.5	13.162	6.326	19.101	1.788	3.644
Ukraine	3.2	37.722	14.219	107.532	3.366	25.595
Slovak Republic	3.2	0.000	22.454	184.551	1.319	18.291
Sri Lanka	4.3	381.381	19.516	48.478	0.828	12.717
Oman	4.9	412.234	9.428	108.592	-3.854	16.483
Dominica	4.5	0.000	15.254	54.002	3.295	15.677
Uzbekistan	3.758	114.335	12.119	42.839	1.601	25.328
Guatemala	3.9	0.456	19.581	51.333	1.843	9.244
Kenya	4.2	286.415	11.363	44.807	2.267	12.816
Myanmar	2.6	3.366	20.672	47.317	6.524	9.856
Ethiopia	3.2	1228.738	4.081	37.173	3.522	14.589
Luxembourg	4.7	6309.473	5.325	438.157	43.303	293.77
Belarus	3.2	1.807	25.934	119.850	3.026	20.108
Costa Rica	3.1	9.284	13.037	62.790	5.586	18.631
Uruguay	4.7	1.842	15.186	44.899	2.562	10.603
Azerbaijan	4.3	41.955	5.768	72.639	7.630	24.727
Panama	6.3	121.567	5.772	91.719	11.193	31.694

Bulgaria	3.9	1.583	14.234	128.067	3.686	25.650
Croatia	4.5	0.775	14.724	96.017	0.326	33.237
Lebanon	3.9	53.902	9.139	121.930	4.974	62.399
Tanzania	3.4	2.337	5.626	47.965	4.297	14.100
Tunisia	3.6	10.354	16.855	92.573	2.245	17.087
Slovenia	5	1.349	23.229	146.754	3.929	25.995
Lithuania	4.9	0.566	19.341	152.376	2.343	27.374
Yemen, Rep.	2.6	0.000	7.458	32.726	-0.041	9.805
Ghana	3.5	0.845	5.328	99.248	8.503	35.830
Jordan	4	169.105	18.467	98.078	3.398	28.797
Serbia	2.7	2.748	19.100	103.100	6.311	23.349
Congo, Dem. Rep.	2.4	0.086	19.035	64.464	-1.441	6.241
Uganda	2.5	0.023	9.199	46.475	3.841	17.832
Paraguay	3.1	1.642	12.007	84.455	1.164	7.577
Latvia	5.2	2.278	12.513	119.091	2.827	25.933
El Salvador	4	13.874	20.912	67.989	2.006	14.987
Trinidad	4	43.198	5.976	84.219	6.870	43.649
Estonia	5.5	0.870	15.837	154.423	-2.903	43.417
Nepal	1.6	4.536	6.297	53.338	0.245	12.415
Zambia	2.2	79.093	7.930	84.316	7.482	9.190
Honduras	4.6	0.502	18.624	107.435	6.448	21.938
Cyprus	4.4	0.231	4.756	122.101	40.957	75.642
Afghanistan	3.3	33.102	12.018	56.979	0.875	11.805
Cambodia	3.7	2.301	17.019	127.864	9.424	33.928
Iceland	6	102.357	12.460	99.879	6.193	42.771

Country Name	Natural gas rents	Oil rents	Mineral rents	GDP	Tax revenue	Tax on income	Inflation, GDP deflator
Ranking Weight	0.02	0.02	0.02	0.08	0.03	0.03	0.06
United States	0.100	0.049	0.069	$18,036,648,000,000.00	9.821	53.542	1.076
China	0.034	0.253	0.510	$11,064,664,793,255.70	10.284	24.903	0.076
Japan	0.001	0.001	0.004	$4,383,076,298,081.86	9.682	51.741	2.035
Germany	0.016	0.011	0.000	$3,363,446,822,668.29	11.559	16.054	1.969
UK	0.106	0.225	0.000	$2,861,090,726,739.55	25.353	33.352	0.550
France	0.000	0.005	0.000	$2,418,835,532,882.33	22.574	25.439	0.638
India	0.073	0.373	0.287	$2,088,841,351,184.16	10.834	44.839	1.901
Italy	0.027	0.045	1.136	$1,821,496,964,400.58	23.632	32.150	0.625
Brazil	0.040	0.858	1.339	$1,803,652,649,613.75	13.651	21.012	7.902
Canada	0.211	0.223	0.543	$1,552,807,652,015.37	11.693	53.043	-0.779
Korea, Rep.	0.000	0.000	0.002	$1,377,873,107,856.33	14.718	24.492	2.210
Russia	3.169	5.579	0.899	$1,365,865,245,098.80	13.996	1.900	8.151
Australia	0.136	0.152	3.832	$1,339,140,527,498.13	21.318	65.334	-0.652
Spain	0.000	0.003	0.034	$1,192,901,186,647.44	12.228	33.963	0.503
Mexico	0.147	1.309	0.625	$1,143,793,184,190.10	9.913	32.889	2.521
Indonesia	0.237	0.586	0.525	$861,933,968,740.33	11.381	34.759	4.231
Netherlands	0.480	0.028	1.136	$750,283,908,173.45	19.608	23.372	0.086
Turkey	0.004	0.029	0.230	$717,879,788,566.76	20.377	17.551	7.429
Switzerland	0.384	1.835	1.136	$670,789,928,809.88	9.496	20.469	-0.562
Saudi Arabia	0.813	22.501	0.058	$646,001,866,666.67	16.327	19.216	-17.191
Argentina	0.197	0.606	0.301	$584,711,485,365.11	9.623	17.041	24.545
Sweden	0.000	1.835	0.194	$495,694,356,611.55	26.069	12.696	2.036
Nigeria	0.328	3.034	0.027	$486,792,837,970.54	1.557	28.319	2.864
Poland	0.073	0.030	0.362	$477,066,454,436.93	15.965	12.705	0.606
Belgium	0.000	1.835	1.136	$455,085,948,763.45	25.710	35.404	0.868
Thailand	0.233	0.409	0.044	$395,168,025,882.03	15.451	36.862	0.222
Norway	2.315	3.043	0.010	$386,578,443,732.56	26.990	31.667	-2.304
Austria	0.026	0.034	0.006	$376,950,249,528.67	26.054	27.010	1.887
Arab Emirates	0.743	11.206	1.136	$370,296,255,956.43	0.365	25.155	-11.213
Egypt	0.774	2.608	0.239	$330,778,550,716.75	12.520	26.155	10.941
South Africa	0.012	0.003	2.192	$314,571,945,857.40	25.587	49.239	3.957
Hong Kong	0.384	1.835	1.136	$309,234,500,374.11	12.500	36.185	3.626
Denmark	0.116	0.383	0.001	$301,307,828,843.61	33.446	41.867	0.904
Malaysia	0.626	1.430	0.203	$296,283,190,372.55	15.613	52.041	-0.371
Singapore	0.384	1.835	1.136	$292,739,307,535.64	13.854	34.716	1.640
Philippines	0.033	0.038	1.247	$292,451,392,606.61	12.888	42.082	-0.632

Colombia	0.119	2.181	0.621	$292,080,155,633.31	13.272	19.730	2.560
Ireland	0.004	1.835	0.029	$283,703,217,034.49	22.445	37.331	4.879
Pakistan	0.584	0.310	0.021	$271,049,886,672.73	10.099	28.849	4.319
Chile	0.016	0.017	11.57	$240,796,388,428.74	19.130	30.361	4.332
Finland	0.384	1.835	0.258	$232,351,114,561.38	20.261	14.738	1.692
Portugal	0.384	1.835	0.124	$199,112,621,714.54	20.883	20.775	2.104
Bangladesh	0.717	1.835	1.136	$195,078,665,827.57	9.025	23.234	5.873
Greece	0.000	0.005	0.054	$194,851,319,174.89	23.604	18.783	-1.044
Vietnam	0.150	1.303	0.189	$193,599,379,094.86	18.977	37.902	-0.191
Peru	0.263	0.377	5.434	$189,111,139,010.08	16.547	33.959	2.238
Czech Republic	0.003	0.010	1.136	$185,156,359,571.12	13.409	14.498	1.002
Kazakhstan	0.843	5.632	2.213	$184,388,432,148.72	13.312	33.616	1.824
Iraq	0.028	28.605	0.003	$180,068,537,409.15	0.914	30.616	-23.495
Romania	0.442	0.335	0.017	$177,954,489,851.96	17.884	17.245	2.919
New Zealand	0.078	0.140	0.235	$173,754,075,210.52	28.420	36.311	-0.115
Algeria	2.589	9.130	0.056	$164,779,467,702.95	37.186	28.241	-7.264
Qatar	5.439	5.852	1.136	$164,641,483,516.48	14.655	40.241	-22.902
Hungary	0.106	0.074	0.000	$121,715,203,207.65	22.988	15.644	1.732
Kuwait	0.648	38.483	1.136	$114,041,209,704.22	1.545	35.241	-27.206
Angola	0.023	10.734	1.136	$102,626,929,544.81	16.456	31.919	-4.018
Morocco	0.004	0.001	2.258	$100,593,283,696.73	23.297	25.396	1.749
Ecuador	0.023	3.283	0.182	$100,176,808,000.00	16.327	25.410	-2.222
Sudan	1.369	0.671	2.192	$97,156,119,150.00	16.327	25.410	17.904
Ukraine	1.369	0.288	1.461	$90,615,023,323.74	18.317	11.808	38.396
Slovak Republic	0.008	0.001	0.018	$87,263,622,047.24	11.926	18.996	-0.216
Sri Lanka	0.384	1.835	0.003	$82,316,172,384.33	10.408	16.180	2.143
Oman	2.465	20.452	0.011	$69,830,949,284.79	2.601	2.393	-18.437
Dominican	0.384	1.835	1.914	$68,102,618,092.10	13.036	26.087	0.884
Uzbekistan	4.587	0.130	4.642	$66,732,736,498.20	16.327	25.410	9.531
Guatemala	0.384	0.094	0.733	$63,794,152,886.04	10.836	29.429	3.266
Kenya	0.384	1.835	0.076	$63,398,041,540.37	15.878	40.892	9.143
Myanmar	0.852	0.149	0.283	$62,600,906,116.10	16.327	25.410	3.939
Ethiopia	0.384	1.835	0.547	$61,539,711,686.69	9.209	16.027	6.360
Luxembourg	0.384	1.835	0.046	$56,799,626,261.51	25.813	28.747	0.398
Belarus	0.005	0.256	1.136	$54,608,962,634.99	15.140	3.453	16.294
Costa Rica	0.384	1.835	0.017	$54,136,834,090.87	13.319	15.071	3.313
Uruguay	0.384	1.835	0.093	$53,442,697,568.72	18.822	18.374	8.698
Azerbaijan	2.212	11.040	0.097	$53,047,140,347.45	12.956	13.592	-8.846
Panama	0.384	1.835	0.070	$52,132,289,747.17	16.327	25.410	0.242
Bulgaria	0.027	0.010	1.497	$50,199,117,547.04	18.496	14.624	2.206
Croatia	0.270	0.116	0.000	$48,732,003,674.38	19.577	7.943	0.116

Lebanon	0.384	1.835	1.136	$47,084,703,150.91	15.484	19.011	1.639
Tanzania	0.081	1.835	2.226	$45,628,247,290.46	12.736	25.822	6.565
Tunisia	0.300	1.813	0.535	$43,015,089,722.68	21.128	26.692	3.380
Slovenia	0.001	1.835	1.136	$42,774,769,768.22	17.964	9.662	0.976
Lithuania	0.384	0.026	1.136	$41,400,137,850.73	4.506	14.944	0.241
Yemen, Rep.	0.348	1.893	1.136	$37,733,919,936.25	9.282	24.345	21.398
Ghana	0.384	1.742	5.544	$37,543,361,203.56	14.866	24.683	17.800
Jordan	0.016	0.001	1.131	$37,517,410,281.69	15.258	13.618	2.281
Serbia	0.041	0.083	0.515	$37,160,332,465.16	19.721	7.596	2.676
Congo	0.002	0.255	14.43	$35,237,742,278.11	8.775	11.879	0.643
Uganda	0.384	1.835	0.006	$27,529,249,701.15	10.861	30.642	5.080
Paraguay	0.384	1.835	1.136	$27,093,938,619.33	12.794	11.614	-0.605
Latvia	0.384	1.835	1.136	$27,002,832,427.64	20.343	9.190	0.388
El Salvador	0.384	1.835	1.136	$25,850,200,000.00	14.533	23.031	0.705
Trinidad	5.313	2.157	1.136	$23,559,287,483.93	26.373	53.004	-9.920
Estonia	0.384	0.109	1.136	$22,459,443,273.82	1.274	18.210	1.038
Nepal	0.384	1.835	1.136	$21,194,888,047.83	13.862	17.997	5.074
Zambia	0.384	1.835	10.08	$21,154,394,545.90	16.148	47.966	6.660
Honduras	0.384	1.835	0.601	$20,420,967,148.94	14.747	21.618	6.300
Cyprus	0.384	1.835	0.062	$19,559,942,331.15	23.267	26.462	-1.261
Afghanistan	0.118	0.038	0.008	$19,331,286,549.33	7.472	3.724	3.817
Cambodia	0.384	1.835	1.136	$18,049,954,289.42	11.095	12.159	1.263
Iceland	0.384	1.835	1.136	$16,779,598,787.16	22.202	27.429	5.928

CURRICULUM VITAE

ALI SAEED AL-ARJANI

EDUCATIONAL BACKGROUND

09/2013 - 2017 PhD Student in Industrial Engineering, University of Wisconsin-Milwaukee.
Title of PhD thesis: "INTRODUCTION OF SIMILARITY COEFFICIENT-BASED CLUSTERING ALGORITHMS TO GLOBAL PETROCHEMICAL FACILITY LOCATION"

Advisors: *Nidal Abu Zahra and Hamid Seifoddini*

02/2011 - 05/2013 M.Sc. in Industrial Engineering, California State University, East Bay
http://www.csueb.edu
Title of Research: "*Organizational structure of Saudi Aramco Information Technology Co. (SAIT).*"

Advisor: *Dr. Hongwei Du*.

09/2003 - 02/2009 B.Sc. in Mechanical Engineering, King Fahad University of petroleum and minerals.
http://www.kfupm.edu.sa
Title of senior Project: "*Simulation of composite material 0-90 layers by ANSYS.*"

Advisor: *Dr. Khalid Al-Dheylan* – May Allah bless him with his jannah

RESEARCH INTERESTS

- Lean Manufacturing
- Group Technology related
- Simulation

PROFESSIONAL EXPRIENCE

06/2012 - 08/2013 Teaching as lecturer in Industrial Engineering Department at Salman Bin Abdul-Aziz University
Courses in English:

- *Engineering Economy.*
- *Engineering Graphic.*
- *Static of Engineering Mechanics.*

06/2009 - 01/2010 Plant Engineer in Saudi Arabian Basic Industries Corporation (SABIC)
I was planner for hot strip plant and scheduling the spare parts for
The down cooler at the end of the production line, which was my area.

HONARS & AWARDS

- Grant of King Abdullah Scholarship to pursue my Master in Industrial Engineering at California State University East Bay Hayward in 2010.
- Passed the Qualification Exam of Saudi Higher Ministry of Education on May 2012 .
- Grant of King Abdullah Scholarship to pursue my PhD Industrial Engineering at UW Milwaukee in 2013.

ACADEMIC ACTIVITIES & PROJECTS

- California State University, East Bay. Research of the organizational structure of Saudi Aramco Information Technology Co. (SAIT). ARAMCO is the biggest petrochemical company in Middle East. Hayward 2012
- Facility Layout Project for new cooling machine in a Pin Cell in JoyGlobal Co. Milwaukee 2013.

COMPUTATIONAL SKILLS

- Graph and drawing softwares. (Auto CAD and Solidwork)
- Simulation software. (ANSYS and Promodel)

REFERENCES

- Saeid Motavalli, Professor, Director of Engineering, Department of Industrial Engineering, California State University, East Bay, Hayward, CA .E-Mail: saeid.motavalli@csueastbay.edu Phone:+1 510-885-2654.
- Farnaz Ganjeizadeh, Associate Professor, Department of Industrial Engineering, California State University, East Bay, Hayward, CA .E-Mail: farnaz.ganjeizadeh@csueastbay.edu Phone:+1 (510) 885-4862.
- Helen Zong, Professor, Department of Industrial Engineering, California State University, East Bay, Hayward, CA .E-Mail: farnaz.ganjeizadeh@csueastbay.edu Phone:+1 (510) 885-4482.