

**CLASSIFICATION OF UNIVERSITIES IN MALAYSIA**

**BOON KOK SIONG**

**UNIVERSITI TEKNOLOGI MALAYSIA**

CLASSIFICATION OF UNIVERSITIES IN MALAYSIA

BOON KOK SIONG

A thesis submitted in partial fulfillment of  
the requirements for the award of the degree of  
Master of Science ( Mathematics)

Faculty of Science  
Universiti Teknologi Malaysia

**DISEMBER 2010**

*Special dedicated to  
My beloved family and friends*

## ACKNOWLEDGEMENT

First of all, I would like to express my sincere appreciation to my supervisor, Assoc. Prof. P.M Dr Ismail Mohamad for his encouragement and guidance as well as suggestions.

Furthermore, I am grateful to Universiti Teknologi Malaysia for funding my master study, the librarians at Sultanah Zanariah Library (PSZ), UTM plays an important role to supply the relevant literatures and resources used in this research.

I would also like to say thank you to all my friends and colleagues that had help me to complete my thesis.

## ABSTRACT

Universities attractiveness are measured based on its research output performance. The objective of this research is to classify 20 public universities in Malaysia in several clusters by using average linkage, K-means clustering methods and correspondence analysis. The source data was found in *Buku Perangkaan 2009* from Ministry of Higher Education (MOHE) and MyMohes system. The classification of these universities are carried out using average linkage and K-means clustering and correspondence analysis is divided into 4 clusters. For clustering method, the distance measured is referred as Euclidean distance whereas for Chi-square distance, it is applied in correspondence analysis. The analyzed results showed the average linkage clustering and K-means clustering are performed at the same cluster solution that are 4 cluster solutions. Besides that, the cluster solution for correspondence analysis method is 4 and for MOHE method showed 3 cluster solutions. Clustering method, correspondence analysis method and MOHE method show different membership in each cluster.

## ABSTRAK

Daya tarikan sesebuah universiti diukur berdasarkan prestasi kajian. Tujuan penyelidikan ini adalah untuk mengklasifikasikan 20 buah institusi pengajian tinggi di Malaysia dalam beberapa kluster dengan menggunakan purata linkage (average linkage), analisis K-mins dan analisis korespondensi. Sumber data adalah diperolehi daripada Buku Perangkaan 2009, Kementerian Pengajian Tinggi dan sistem MyMohes. Dengan menggunakan purata linkage, analisis K-mins dan analisis korespondensi, universiti-universiti dapat diklasifikasikan kepada 4 kluster. Untuk kaedah pelaksanaan kluster, jarak diukur adalah berdasarkan jarak Euclidean manakala untuk jarak *Chi-square*, ia diaplikasikan dalam analisis korespondensi. Keputusan dianalisis dalam kajian ini menunjukkan bahawa keputusan yang dihasilkan oleh purata linkage dan analisis K-mins adalah sama iaitu penyelesaian 4 kluster. Selain itu, penyelesaian kluster untuk kaedah analisis korespondensi adalah 4 kluster dan untuk kaedah Mohe menunjukkan 3 penyelesaian kluster. Kaedah pelaksanaan kluster, kaedah analisis korespondensi dan kaedah MOHE menunjukkan keahlian yang berbeza dalam setiap kluster.

## TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
	<b>DECLARATION</b>	<b>ii</b>
	<b>DEDICATION</b>	<b>iii</b>
	<b>ACKNOWLEDGEMENT</b>	<b>iv</b>
	<b>ABSTRACT</b>	<b>v</b>
	<b>ABSTRAK</b>	<b>vi</b>
	<b>TABLE OF CONTENTS</b>	<b>vii</b>
	<b>LIST OF TABLES</b>	<b>x</b>
	<b>LIST OF FIGURES</b>	<b>xii</b>
	<b>LIST OF ABBREVIATIONS</b>	<b>xiii</b>
	<b>LIST OF APPENDICES</b>	<b>xvi</b>
<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
	1.1 Background of The Study	1
	1.2 Statement Of Problem	4
	1.3 The Objectives of the Study	4
	1.4 Scope of the Study	5
	1.5 Significance of the Study	5
	1.6 Limitation of the Research	5
	1.7 Structure of Dissertation	6
<b>2</b>	<b>LITERATURE REVIEW</b>	<b>7</b>
	2.0 Introduction	7

2.1	Distance Measures	8
2.2	Cluster Analysis	10
	2.2.1 Clustering Techniques	11
	2.2.1.1 Hierarchical Cluster Analysis	12
	2.2.1.2 Non-hierarchical Clustering Methods (NHCA)	14
2.3	Correspondence Analysis	16
<b>3</b>	<b>Data Exploration</b>	19
	3.1 Introduction of the Data	19
	3.2 Exploration of Data	21
	3.3 Analyzing Data	26
	3.4 Standardization of Data	31
	3.5 Summary	31
<b>4</b>	<b>METHODOLOGY</b>	35
	4.0 Introduction	35
	4.1 Hierarchical Cluster Analysis Methods	35
	4.1.1 Single Linkage Clustering	37
	4.1.2 Complete Linkage Clustering	38
	4.1.3 Average Linkage Clustering	39
	4.1.4 Centroid Method	40
	4.1.5 Ward's Method	40
	4.2 K-means Clustering	42
	4.3 Correspondence Analysis	43
	4.3.1 Chi-squared Distance	46
	4.3.1.1 Row Distance	46
	4.3.1.2 Column Distance	47
	4.3.2 Inertia	48
	4.3.3 Singular Value Decomposition (SVD)	49
	4.4 Summary	51

<b>5</b>	<b>RESULTS AND DATA ANALYSIS</b>	<b>52</b>
5.0	Introduction	52
5.1	Average Linkage Clustering	52
5.2	K-means Clustering	57
5.3	Correspondence Analysis	63
5.4	Comparison Between the Data	81
5.5	Summary	84
<b>6</b>	<b>CONCLUSIONS AND RECOMMENDATIONS</b>	<b>85</b>
6.0	Introduction	85
6.1	Conclusions	85
6.2	Recommendation	86
	<b>REFERENCES</b>	<b>87</b>
	<b>APPENDIX A</b>	<b>89</b>

## LIST OF TABLES

<b>TABLE NO</b>	<b>TITLE</b>	<b>PAGE</b>
1.1	List of Public Universities in Malaysia.	3
3.1	Criteria Of Public Universities	20
3.2(a)	The Criteria of Public Universities in Research	22
3.2(b)	The Criteria of Public Universities in Academic Staffs (Foreign and Local Nationality)	23
3.2(c)	The Criteria of Public Universities in Qualification of Academic Staffs Level (Foreign and Local Nationality)	24
3.2(d)	The Criteria of Public Universities in Enrolment Students (Foreign and Local Nationality)	25
3.3	Mean and Standard Deviation of Each Criteria of Public Universities	32
3.4(a)	The Standardized Data for Criteria : Research And Academic Staffs	33
3.4(b)	The Standardized Data for Criteria: Qualification Level of Academic Staffs and Enrolment Student ( foreign and Local Nationality)	34
4.1	Contingency Table	44
5.1	Agglomeration Schedule	55
5.2	Cluster Membership Obtained by Average Linkage Clustering	57
5.3	Initial Cluster Centers	58
5.4	Iteration History	59
5.5	Cluster Membership	59
5.6	Final Cluster Centers	60
5.7	Number Of Cases in Each Cluster	61
5.8	The Cluster Membership of K-Means Clustering	61

5.9(a)	Contingency Table	64
5.9(b)	Contingency Table	65
5.9(c)	Contingency Table	66
5.10(a)	Row Profile	67
5.10(b)	Row Profile	68
5.10(c)	Row Profile	69
5.11(a)	Column Profile	70
5.11(b)	Column Profile	71
5.11(c)	Column Profile	72
5.12	Inertia and Chi-square Decomposition of the Correspondence Analysis	73
5.13	Principal Coordinate of Row and Column Profile	74
5.14	The classification of Quality	78
5.15	Result From Correspondence Analysis	81
5.16	Comparison the Result Between Average Linkage, K-means Clustering, Correspondence Analysis and MOHE	83

## LIST OF FIGURES

<b>FIGURE NO</b>	<b>TITLE</b>	<b>PAGE</b>
2.1	Main Idea of Hierarchical Techniques	13
2.2	Basic K-means Algorithm	15
3.1	The Criteria of Public Universities in Research	26
3.2	The Criteria of Public Universities in Academic Staffs (Foreign and Local Nationality)	27
3.3	The Criteria of Public Universities in Qualification of Academic Staffs Level (Foreign and Local Nationality)	29
3.4	The Criteria of Public Universities in Enrolment Students (Foreign and Local Nationality)	30
4.1	Cluster distance, single linkage method	37
4.2	Cluster distance, complete linkage method	38
4.3	Cluster distance, average linkage method	39
4.4	Cluster distance, Centroid method	40
4.5	Cluster distance, Ward's method	41
5.1	Dendrogram using Average Linkage (Between Groups)	56
5.2	The symmetric map between Public Universities and the Criteria of the public universities	77
5.3	Quality and Quantity of Researcher, Qualification of Staff and Postgraduate	78

## LIST OF ABBREVIATIONS

EDEIN	- Enrollment of Degree Students 2009 (International)
EDEL	- Enrollment of Degree Students 2009 (Local)
EMAIN	- Enrollment of Master Students 2009 (International)
EMAL	- Enrollment of Master Students 2009 (Local)
EPHDIN	- Enrollment of PhD Students 2009 (International)
EPHDL	- Enrollment of PhD Students 2009 (Local)
ESS	- Error Sum Of Squares
HCA	- Hierarchical Cluster Analysis
IPTA	- Higher Learning Institution
LECIN	- Number of Lectures (International)
LECL	- Number of Lectures (Local)
MOHE	- Ministry of Higher Education
NOC	- Number Of Commercial 2009
NOPB	- Number of Publication of Book 2009
PMIN	- Number of Professor Madya (International)
PML	- Number of Professor Madya (Local)
PROFIN	- Number of Professors (International)
PROFL	- Number of Professors (Local)
QDEIN	- Qualification of Lecturer at Degree Level (International)
QDEL	- Qualification of Lecture at Degree Level (Local)
QMALIN	- Qualification of Lecturer at Master Level (International)
QMAL	- Qualification of Lecturer at Master Level (Local)
QPHDIN	- Qualification of Lecturer at PhD Level (International)
QPHDL	- Qualification of Lecturer at PhD Level (Local)
QS	- Quacquarelli Symonds
RGS	- Research Grant Started 2009
RGF	- Research Grant Ended 2009

THE	- Times Higher Education
UM	- Universiti Malaya
UDM	- Universiti Darul Iman Malaysia
UIAM	- Universiti Islam Antarabangsa Malaysia
UiTM	- Universiti Teknologi MARA
UKM	- Universiti Kebangsaan Malaysia
UMK	- Universiti Malaysia Kelantan
UMP	- Universiti Malaysia Pahang
UMS	- Universiti Malaysia Sabah
UMT	- Universiti Malaysia Terengganu
UniMAP	- Universiti Malaysia Perlis
UNIMAS	- Universiti Malaysia Sarawak
UPM	- Universiti Putra Malaysia
UPSI	- Universiti Pendidikan Sultan Idris
UPNM	- Universiti Pertahanan Nasional Malaysia
USM	- Universiti Sains Malaysia
USIM	- Universiti Sains Islam Malaysia
UTM	- Universiti Teknologi Malaysia
UTeM	- Universiti Teknikal Malaysia Melaka
UTHM	- Universiti Tun Hussein Onn Malaysia
UUM	- Universiti Utara Malaysia

## LIST OF APPENDICES

<b>APPENDIX</b>		<b>PAGE</b>
<b>A</b>	(i) The Number of University Academic Staff by Position and Gender, in 2009 (Local)	89
	(ii) The Number of University Academic Staff by Position and Gender, in 2009 (International )	90
	(iii) The Number of University Academic Staff by Educational level and gender, in 2009 (Local)	91
	(iv) The Number of University Academic Staff by Educational level and gender, in 2009 (International)	92
	(v) University Enrollment by Level of Education and Gender, year 2009 (Local)	93
	(vi) University Enrollment by Level of Education and Gender, year 2009 (International)	94
	(vii) Statistical research for 20 public universities in 2008 and 2009	95

## **CHAPTER 1**

### **INTRODUCTION**

#### **1.1 Background of The Study**

In our lives, education plays an important role. Since the day we were born, everyone needs to be educated one way or another. There is a rapidly growing demand for higher education in the world today. Although it is difficult to achieve a higher education, the rewards that comes with it such as for self-improvement, job assurance, character development and social standings are the things that many of us seek to satisfy ourselves. Surely it is hard to go to school a bit longer, but learning for personal knowledge would greatly improve us. We get to learn skills for problem solving which would teach us how to figure things out by ourselves. A better education will also gives us a lot more experience in things that we might like or otherwise. Undeniably, knowledge is a very powerful thing that can change the lives of others besides ours.

These days, it is becoming a necessity for parents to send their children to study up to the university level. Getting a higher education is a good and welcoming trend in a rapidly developing society. Yet there are two different points of views over this higher education issue. While some people strongly believe that having a degree has numerous benefits, there are those who opine differently.

Undoubtedly, having university qualifications is beneficial for the advancement of a country; meaning that it enhances the quality of its citizen and in turn making the country more competitive in various fields. Besides gaining a lot more knowledge and skills from the universities courses, the students improve their thinking skills which should booster their competitiveness in the society. Furthermore, universities also offer all kinds of lessons to teach students how to distinguish right from wrong and on how to deal with unforeseen circumstances that just might poop up. Therefore, they would not be misled nor be misinformed and fall into the trappings and temptations of a society. Again it can be argued that the universities have their merits and numerous advantages for a country and its citizens.

A public university is a university that is predominantly funded by public through a national or sub national government, as opposed to private universities. The List of Public Universities in Malaysia is shown in Table 1.1 :

The Public universities (Government's Universities) produce graduates who can acquire and apply their knowledge for the benefit of the Malaysian society. They are run by the Malaysian Government guided by objectives which are in line with the Ministry of Higher Education (MOHE) policies. The MOHE's main objective is to produce a human capital that is dynamic, competitive, and foremost equipped with a first class mentality to face the challenges of a changing world. The 20 public universities of Malaysian Government are grouped into three categories namely the research universities, Comprehensive universities and Focused universities. The five research universities are Universiti Malaya (UM) (2006), Universiti Putra Malaysia (UPM)(2006), Universiti Kebangsaan Malaysia (UKM)(2006), Universiti Sains Malaysia (USM)(2008) and Universiti Teknologi Malaysia (UTM) (2010).

Since the year 2006, the Ministry of Higher Education (MOHE) had chosen 4 designated research universities and each of the universities was awarded RM100 million grant to do research work, development and commercialization activities. Such action helped in gaining popularity for Malaysia as an aspiring education hub in the world.

**Table 1.1** : List of Public Universities in Malaysia.  
[www.etawau.com/edu/IndexUniversityGovernment.htm](http://www.etawau.com/edu/IndexUniversityGovernment.htm)

Public universities in Malaysia	Date of Operation
Universiti Malaya (UM)	7-1-1962
Universiti Sains Malaysia (USM)	1969
Universiti Kebangsaan Malaysia (UKM)	18-5-1970
Universiti Putra Malaysia (UPM)	4-10-1971
Universiti Teknologi Malaysia (UTM)	1-4-1975
Universiti Teknologi MARA (UiTM)	26-8-1999
Universiti Islam Antarabangsa Malaysia (UIAM)	10-5-1983
Universiti Utara Malaysia (UUM)	16-2-1984
Universiti Malaysia Sarawak (UNIMAS)	24-12-1992
Universiti Malaysia Sabah (UMS)	24-11-1994
Universiti Pendidikan Sultan Idris (UPSI)	24-2-1997
Universiti Sains Islam Malaysia (USIM)	13-3-1998
Universiti Malaysia Terengganu (UMT)	15-7-1999
Universiti Tun Hussein Onn Malaysia (UTHM)	30-9-2000
Universiti Teknikal Malaysia Melaka (UTeM)	1-12-2000
Universiti Malaysia Pahang (UMP)	16-2-2002
Universiti Malaysia Perlis (UniMAP)	2-5-2002
Universiti Darul Iman Malaysia (UDM)	1-1-2006
Universiti Malaysia Kelantan (UMK)	14-6-2006
Universiti Pertahanan Nasional Malaysia (UPNM)	10-11-2006

## **1.2 Statement Of Problem**

According to W.Y.Aw (2009), a university's attractiveness can be determined by many factors such as the internal and external factors. The internal factors refer to the university characteristics like university's facilities, teaching services, academic staffs, and programs offered while external the factors are things related to the socio-economic conditions and political stability. Another aspect on measuring the university's attractiveness normally and widely used by the international institutions and scholars is the university's research output performance. Universities and faculties are ranked based on their research outputs or publications. Various methods such as the THE – QS World University Rankings and Ranking Web are amongst the many research assessments used by scholars both in the past and present in ranking the faculties and universities.

The 20 public universities in Malaysia are classified into the three categories namely the research universities, comprehensive universities and focused universities. The classification is based on the internal and external factor defined by MOHE. The cluster analysis is a method of classification of items whereas correspondence analysis is a multivariate method of analysis specializing in contingency table. Correspondence analysis method displays the important information into geometric view and allows us to observe the similarities and dissimilarities among the items according their positions. This study compare the classification of these universities by clustering method, correspondence analysis method and the MOHE classification method.

## **1.3 The Objectives of the Study**

The main objectives of this research are:

- (i) To classify public universities in Malaysia using two multivariate techniques: Hierarchical Clustering and Correspondence Analysis.

- (ii) To compare the result of categorization of public universities based on hierarchical clustering and correspondence analysis.
- (iii) To compare the MOHE categorization of universities with the results obtained by hierarchical clustering and correspondence analysis.

#### **1.4 Scope of the Study**

This study will look into the concepts in the hierarchical cluster analysis namely average linkage, K-means clustering and Correspondence Analysis while understanding its application by using its methods in real problems. Besides that, this study will discover how to classify the variables according to their characteristics. The sample of this study is the 20 public universities in Malaysia.

#### **1.5 Significance of the Study**

Multivariate techniques provide an alternative approach to classify 20 public universities in Malaysia whereas correspondence analysis displays a graphical representation of the relationships between the 20 public universities and the criteria of 20 public universities into geometric space. The findings of this study will help the local and foreign students to identifying the similarity and differences of the characteristics between these 20 public universities in Malaysia.

#### **1.6 Limitation of the Research**

This research only focused on four main criteria provided by MyMohe system and *Buku Perangkaan 2009* like the research grants, the number of international and local academic staffs, the qualification of international and local academic staff and the enrolment of international and local postgraduates .

## 1.7 Structure of Dissertation

This dissertation is organized into six chapters. The first chapter presents the background, the statement of problem, the objectives and the scope of the research.

Chapter Two is the literature review. This chapter is focused on the clustering method namely hierarchical clustering and non hierarchical clustering and correspondence analysis due to their common characteristics of grouping or clustering similar items or variables into an unknown number of groups or clusters based on distance measures

Chapter three discussed the descriptions of data set which is used in this study. Chapter Four is focused on clustering methods and correspondence analysis method. Clustering methods are generally classified as hierarchical clustering and non-hierarchical clustering. The two main categories of methods for hierarchical cluster analysis are divisive methods and agglomerative methods. Correspondence analysis is a descriptive technique to analyze simple two way and multi-way tables containing some measure of correspondence between the rows and columns. Its goal is to observe the similarities and dissimilarities items from a contingency table.

Chapter Five discusses the result obtained by average linkage and K-means Clustering method and Correspondence Analysis method. The SPSS software is used to obtain the result from clustering methods whereas the Matlab software is used to obtain the result from Correspondence Analysis Method. The MyMohes system and *Buku Perangkaan 2009* data is used as the data source.

Finally, Chapter Six concludes the study. This chapter summarizes the study and conclusions based on analysis and the results of the study. Suggestions for further research are also recommended in this chapter.

## REFERENCES

- Alan J. I (2008). *Modern Multivariate Statistical Techniques : Regression, Classification, and Mananifold Learning*. Department of Statistics Temple University.
- Batcher, J (2002). *Cluster analysis*. University Erlangen-Nuremberg.
- Bendixen, M (1996)., *A Practical Guide to the Use of Correspondence Analysis in Marketing Research*. (electronic version). Marketing Research On-Line, 2003, pg16-38
- Emmett J. Ientilucci, Chester F. Carlson (2003) “*Using the Singular Value Decomposition*” Center for Imaging Science, Rochester Institute of Technology
- Greenacre, M. J.(1984) *Theory and Applications of Correspondence Analysis*. Academic Press. London.
- Greenacre M.J (2007). *Correspondence Analysis in Practice*. Chapman & Hall/CRC, Boca Raton, FL, 2nd edition
- G.J Gan, C.Q.Ma and J.H. Wu (2007). *Data Clustering :Theory, Algorithms, and Applications* American Statistical Association.
- Hair, J.F. et al. (1992) *Multivariate Data Analysis* (3rd ed.). New York: Macmillan.
- ICF Consulting. (2003). *The Role of University Today: Critical Partners in Economic Development and Global Competitiveness*

Lebart, L., Morineau, A., and Warwick, K.M. (1984), *Multivariate Descriptive Statistical Analysis: Correspondence Analysis and Related Techniques for Large Matrices*, New York: John Wiley & Sons, Inc.

List of Government Universities Malaysia :  
[www.etawau.com/edu/IndexUniversityGovernment.htm](http://www.etawau.com/edu/IndexUniversityGovernment.htm)

MacQueen, J. B. (1967). Some Methods For Classification and Analysis of Multivariate Observations. Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability. University of California Press. Pp 281-297.

MOHE. (2009). Perangkaan Pengajian Tinggi Malaysia Tahun 2009. (Buku Perangkaan 2009). Putrajaya, Malaysia: Bahagian Perancangan dan Penyelidikan, Jabatan Pengajian Tinggi.

Richard A. Johnson, Dean W. Wichern. (2007). *Applied Multivariate Statistical Analysis* 6<sup>th</sup> edi. United State of America.

Rui Xu and Donald C.W (2009). *Clustering*. IEEE Computational Intelligence Society. A John Wiley& Son, Inc, Publication

Tryon, R. C. (1939). *Cluster analysis*. New York: McGraw-Hill

W.Y Aw (2009). *Measuring Research Performance of Selected Public Universities: Implication For International Students Attraction*. Graduate School Of Management, University Putra Malaysia.

Yaves, U and Shem Well, D.J (1996) “Graphical Representation of University Image: A Correspondence Analysis” *Journal of Marketing for Higher Education*, Vol 7, No 2, pg 75-84