Jabatan Perangkaan Malaysia

KOLOKIUM STATISTIK 2017

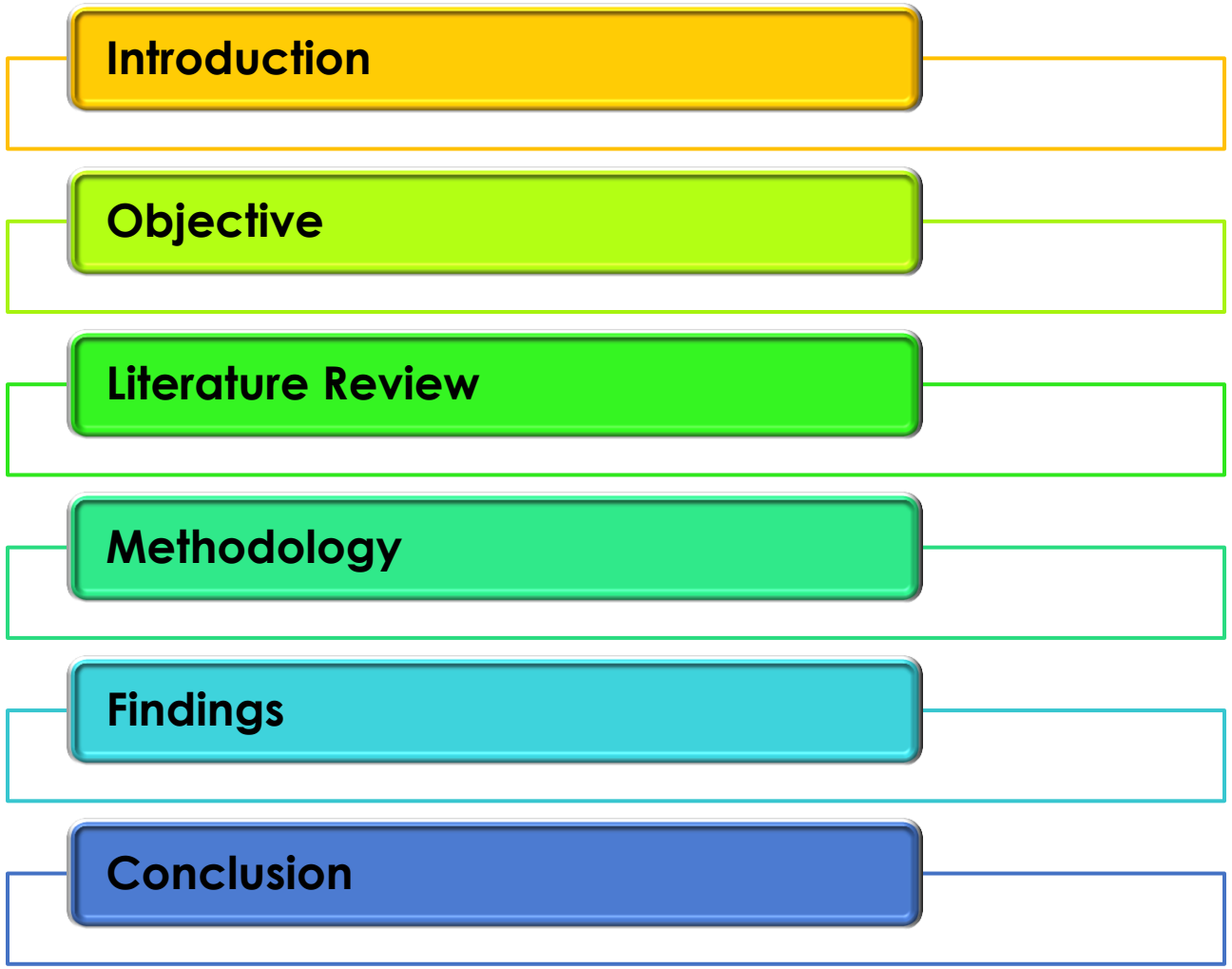
FUNCTIONAL ANALYSIS OF INDUSTRIAL CLUSTERS IN MALAYSIA

28 – 29 September2017

"MEMACU GENERASI ANALITIK KE ARAH MENDEPANI FENOMENA STATISTIK GLOBAL"

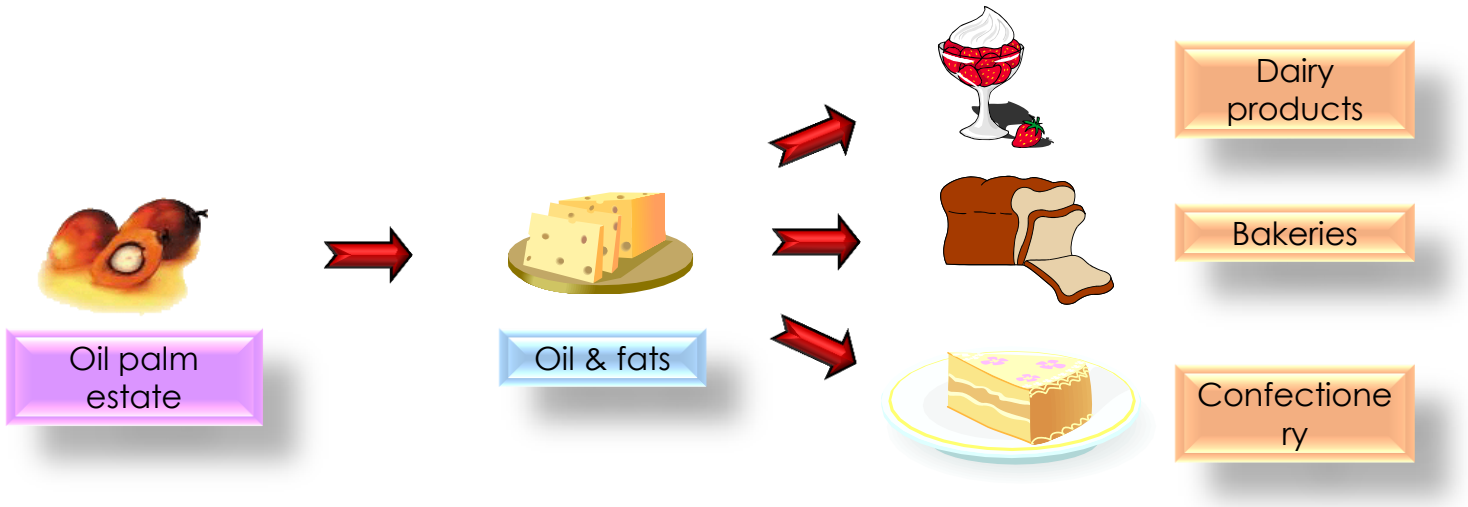
Dr. Norhayati Shuja'
Bahagian Metodologi dan Penyelidikan

Outlines


Introduction
Objective
Literature Review
Methodology
Findings
Conclusion

Introduction

- Industry cluster is a group of industries that has strong complementary linkages.
- Industries that have strong supply and buyer relationship will provide better information of inter-industry dependence.



Introduction

- There have been several empirical studies in the field of industrial clustering.
- However, they differ not only in definitions used and areas studied, but also in the objectives and methodology.
- This study employed Cluster Analysis for grouping industries based on input-output linkages which has been applied for many countries.
- Cluster analysis is useful in analyzing which sectors are strongly connected to each other.

Objective

- The objective of this paper;
  - to identify the national industry clusters in Malaysia by examining the inter-industry selling and purchasing relationship based on the Malaysia's 2010 Input-Output table using cluster analysis approach.

Literature Review

Researcher	Study	Data	Method(s)	Findings
Stan Czamanski (*Papers in Regional Science*, 1971)	i. Developed a technique for identifying industrial clusters with strong internal and weak external flows. ii. Verified the hypothesis that in an urban agglomeration, the ancillary links with supplier based on outputs. iii. Ascertained the minimum size of Input-Output tables usable for sophisticated regional analysis.	• I-O table of the US for 1958 (77 x 77) and 1963 (85 x 85) • The 1959 Philadelphia I-O table (89x89) • The 1963 Washington I-O table (54x54) • The west Virginia (48x48) • The 1960 Nova Scotia (58x58)	Multivariate analysis. Calculate the eigenvalues of the *R* matrix. The *R* matrix is the matrix that consists of the highest of the four correlation coefficients.	Multivariate analysis may be a useful tool for analyzing linkages existing in regional economies.

Literature Review

Researcher	Study	Data	Method(s)	Findings
Joel Bergman, Peter Greenston & Robert Healy (*Journal of Urban Studies*, 1972)	Described set of industrial clusters and discuss alternative methods of defining industrial clusters.	The 1963 employment data of 144 manufacturing activities and 42 non-manufacturing activities.	Factor analysis	The grouping of industries with similar location pattern was identified. The 186 industries resulted in 42 clusters.
Howard Roepke, David Adams & Robert Wiseman (*Journal of Regional Science*, 1974)	Identifying functually related groups of industries. Testing the feasibility of the approach.	Input-output table of 1965 for Canadian Province of Ontario. A 44 by 44 matrix of interindustry product linkages.	Factor analysis, R-mode & Q-mode factoring	The factor analysis of input-output data may prove to be a technique of value in both spatial and aspatial analysis.
Joel Bergman, Peter Greenston & Robert Healy (*Journal of*	Extended their studies using a data base which consist of 480 industrial classification.	The 1965 employment data in each of 480 activities in each of 311 metropolitan areas in US	Hierarchical clustering analysis	The clustering is quite different from the Standard Industrial Classification (SIC), except at the most aggregated level. Revealed 40 groups of industries

Literature Review

Researcher	Study	Data	Method(s)	Findings
Edward J. Feser & Edward M. Bergman (*Journal of Regional Studies*, 2000)	Suggested a means of using detailed information on national interindustry linkages to identify existing and potential clusters in US.	Detailed 1987 benchmark US Input-Output table.	Factor analysis	A national set of benchmark or template technological clusters effectively represent strategically important alignments of underlying detailed sectors.
Edward W. Hill & John F. Brennan (*Economic Development Quarterly*, 2000)	Identified the industries in which a region has its greatest competitive advantage.	194 industries in the Cleveland-Akron (Ohio) consolidated metropolitan area.	Combining cluster and discriminant analysis	5 of 10 industry clusters clearly met the definition of having competitive advantage.
Edward Feser (*Regional Economies Applications Laboratory*, 2005)	Developed an alternative methodology for identifying benchmark value chain clusters	Benchmark Input-Output Accounts of the US, 1997	Czamanski four correlations & Factor analysis	The use of more sophisticated clustering algorithm separates strong and weak linkages with greater precision. The clusters appear to correspond to basic industry.

Methodology

- The identification of linkages and industrial clustering involve four stages.

Stage 1	Calculate input and output coefficient based on I-O transaction matrix, 120 industries by 120 industries
Stage 2	Calculate four pairwise linkage ratios, *r*, for each pair of industries
Stage 3	Produce R_{max} matrix by selecting the maximum value of *R*. Set linkage threshold at 0.02 for purchases and 0.01 for sales.
Stage 4	Identify core cluster by using Ward's hierarchical clustering algorithm

Methodology

Stage 1

- Calculate the input and output coefficients of an individual industry based on I-O transaction matrix, 120 industries by 120 industries.
- Obtain the forward and backward linkages based on input and output coefficient matrix.

$$x_{ij} = \frac{z_{ij}}{P_j} \qquad y_{ij} = \frac{z_{ij}}{S_i}$$

Where,

P_j total purchases by industry j

S_i total sales by industry i

z_{ij} value of purchases by industry j from industry i

x_{ij} ratio of purchases by industry j from industry i to the total purchases by industry j

y_{ij} ratio of sales by industry i to industry j to the total sales by industry i

Methodology

Stage 2

- Calculate four pairwise linkage ratios, r, for each pair of industries.

$r(x_i, x_j)$: industry i and industry j are correlated by having similar purchasing patterns

$r(y_i, y_j)$: industry i and industry j are correlated by having similar selling patterns

$r(x_i, y_j)$: industry i has a purchasing pattern which is similar as industry j's selling pattern

$r(y_i, x_j)$: industry i has a selling pattern which is similar as industry j's purchasing pattern

Methodology

Stage 3

Find the strongest linkages among the purchases and sales coefficients for each pair of industries.

$$I_{ij}^{SS} = S_i \cap S_j, U_{ij}^{SS} = S_i \cap S_j \quad I_{ij}^{BB} = B_i \cap B_j, U_{ij}^{BB} = B_i \cap B_j$$

$$I_{ij}^{SB} = S_i \cap B_j, U_{ij}^{SB} = S_i \cap B_j \quad I_{ij}^{BS} = B_i \cap S_j, U_{ij}^{BS} = B_i \cap S_j$$

I_{ij}^{SS} : industries that are selling to both industry i and industry j

I_{ij}^{BB} : industries that are purchasing from both industry i and industry j

I_{ij}^{SB} : industries that are selling to industry i and buying from industry j

I_{ij}^{BS} : industries that arebuying from industry i and selling to industry j

U_{ij}^{SS} : a group of suppliers to either industry i or industry j, or both

U_{ij}^{BB} : a group of buyers from either industry i or industry j, or both

U_{ij}^{SB} : a group of suppliers to industry i and buyers from industry j

U_{ij}^{BS} : a group of buyers from industry i and suppliers to industry j

Some industries purchase intermediate input from industry i *(j)* and supply output to industry j *(i)*

Methodology

Stage 3
continued

Calculate the *R* measures, a linkage threshold, $\alpha = 0.02$ for purchases and 0.01 for sales are assigned to industry *i* and industry *j* .

$$R_{ij}^{SS} = \frac{I_{ij}^{SS}}{U_{ij}^{SS}}, R_{ij}^{BB} = \frac{I_{ij}^{BB}}{U_{ij}^{BB}}, \quad R_{ij}^{SB} = \frac{I_{ij}^{SB}}{U_{ij}^{SB}}, R_{ij}^{BS} = \frac{I_{ij}^{BS}}{U_{ij}^{BS}}$$

R_{ij}^{SS} : ratio of the number of common suppliers to industries *i* and *j* over the total number of suppliers to industries *i* and

R_{ij}^{BB} : ratio of the number of common buyers to industries *i* and *j* over the total number of buyers to industries *i* and *j*

R_{ij}^{SB} & R_{ij}^{BS} : measures of second-tier of linkages between industries *i* and *j*.

If the ratio of R_{ij}^{SS} (R_{ij}^{BB}) equals to 1 or 0, then industry *i* and *j* have the same supplying (buying) pattern.

Methodology

Stage 3
continued

Selecting the maximum of the four *R* measures produces a new matrix (R_{max} matrix)

$$R_{ij} = R_{ji} = \max\left[R(x_i, x_j), R(y_i, y_j), R(x_i, y_j), R(y_i, x_j)\right]$$

where, the x and y values are inter-industrial

purchases and sales coefficients

respectively

Methodology

Stage 4

Identify core cluster by using Ward's hierarchical clustering algorithm.

- The hierarchical clustering technique using the Ward's method and applying squared Euclidean Distance as the distance or similarity measure to the R_{max} matrix was carried out to identify the clusters.
- Ward's method uses an analysis of variance approach to evaluate the distances between clusters.
- The cluster membership was assessed by calculating the total sum of squared deviations from the mean of a cluster.

Findings

- The table presents part of the agglomeration schedule for the cluster analysis performed for the Malaysia Input-Output Table for 2005 (the final 24 of all 119 stages).
- There are three possible solutions: 12 clusters, 10 clusters and 3 clusters with a large increase in the agglomeration coefficients at the 118th, 111th, and 109th stages.

Stage	No. of Clusters in Solution	Agglomeration Coefficient	Slope [a] Percentage Change in Agglomeration Coefficient	Acceleration [b] Percentage Change in Slope Coefficient
96	24	72.776	2.203	-1.934
97	23	74.359	2.176	-1.238
98	22	75.967	2.162	-.643
99	21	77.630	2.189	1.277
100	20	79.379	2.252	2.863
101	19	81.185	2.275	1.013
102	18	83.091	2.349	3.241
103	17	85.236	2.581	9.894
104	16	87.394	2.532	-1.899
105	15	89.602	2.526	-.250
106	14	92.119	2.810	11.245
107	13	94.818	2.930	4.283
108	**12**	**97.544**	**2.874**	**-1.902**
109	11	100.538	3.070	6.794
110	**10**	**103.580**	**3.026**	**-1.435**
111	9	106.972	3.275	8.257
112	8	111.070	3.831	16.955
113	7	116.004	4.442	15.945
114	6	121.513	4.749	6.927
115	5	128.287	5.575	17.383
116	4	135.993	6.007	7.751
117	**3**	**143.993**	**5.883**	**-2.063**
118	2	155.470	7.970	35.481
119	1	185.260	19.161	140.406

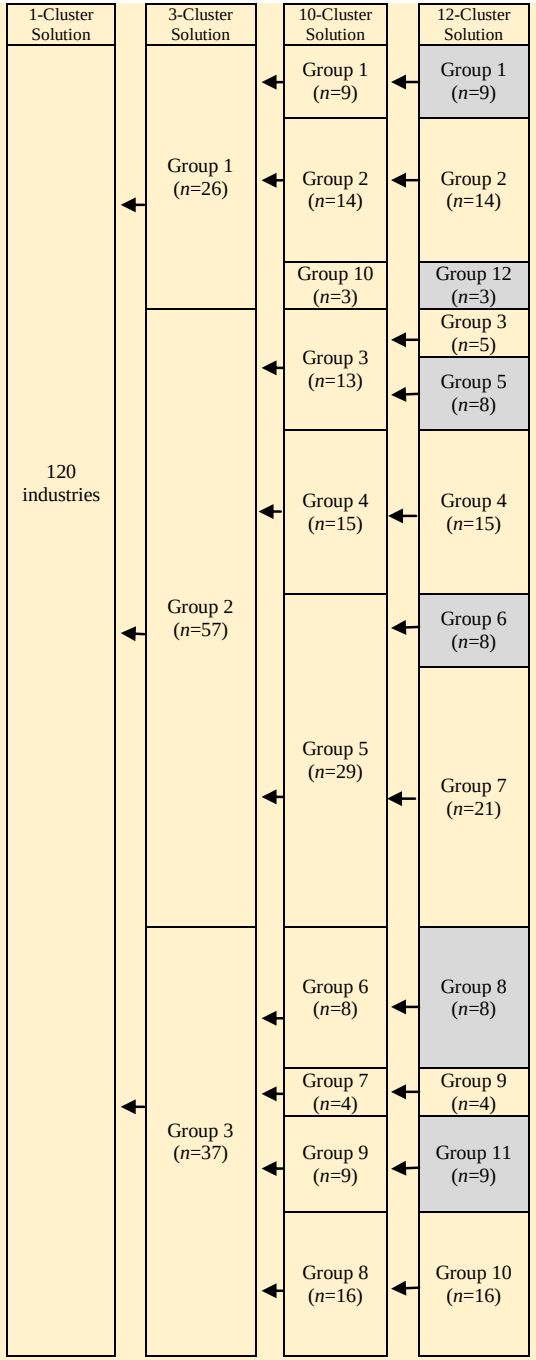
(a) The slope coefficient of the agglomeration schedule is the percentage change in the agglomeration coefficient given in the third column.

(b) Acceleration is the change in the slope of the agglomeration coefficient, calculated as the percentage change in the number contained in the fourth column.

- The percentage change of the slope and acceleration is also high at the three stages.
- Among the three clusters, the 12-cluster is more homogeneous based on the characteristics of the industries.

Findings

- The table shows the 3, 10, and 12-cluster solutions for the 120 industries.
- The heterogeneity of the industry classification depends on the size of the clusters.
- Reading from right to left, the industrial clusters get more and more heterogeneous.
- The homogeneity of the industrial clusters depends on the size of the cluster solution.



Industry		
Paddy	Fruits	Flower Plants
Food Crops	Rubber	Poultry Farming
Vegetables	Oil Palm	Forestry and Logging
Other Agriculture	Fishing	Other Food Processing
Other Livestock	Dairy Production	Wine and Spirit
Meat and Meat Production	Bakery Products	Soft Drink
Leather Industries	Confectionery	Tobacco Products
Soap, Perfumes, Cleaning & Toilet Preparations	Preservation of Fruits and Vegetables	
Highway, Bridge and Tunnel Operation Services	Private Non-Profit Institution Ownership of Dwellings	
Crude Oil and Natural Gas	Other Mining and Quarrying	Stone Clay and Sand Quarrying
Metal Ore Mining Fertilizers	Fertilizers	
Oils and Fats	Motorcycles	Rubber Gloves
Petroleum Refinery	Research and Development	Rubber Products
General Purpose Machinery	Rubber Processing	
Preservation of Seafood	Electrical Machinery and Apparatus	Electric Lamps and Lighting Equipment
Industrial Machinery		
Grain Mills	Measuring, Checking & Industrial Process Equipment	Medical, Surgical and Orthopaedic Appliances
Animal Feeds		
Recycling		Other Electrical Machinery
Watches and Clocks	Optical Instruments and Photographic Equipment	Insulated Wires and Cables
Domestic Appliances		Special Purpose Machinery
Yarn and Cloth	Wearing Apparel	Printing
Finishing of Textiles	Footwear	Tyres
Other Textiles	Publishing	
Sawmilling and Planning of Wood	Other Chemicals Product	Sheet Glass and Glass Products
	Other Transport Equipment	
Veneer, Sheets, Plywood, Laminated & Particle Board	Plastics Products	Basic Precious and Non-Ferrous Metals
	Cement, Lime and Plaster	
Builders' Carpentry and Joinery	Concrete & Other Non-Metallic Mineral Products	Pharmaceuticals, Chemicals & Botanical Product
Iron and Steel Products	Basic Chemicals	
Wooden and Cane Containers	Other Manufacturing	Clay and Ceramic
Other Wood Products	Paints and Varnishes	Structural Metal Products
Paper and Paper Products and Furniture	Castings of Metals	
Office, Accounting and Computing Machinery	Other Fabricated Metal Products	Electricity and Gas Waterworks
Semi-Conductor Devices,Tubes and Circuit Boards	TV, Radio Receivers & Transmitters & Associated Goods	Ships & Boats Building, Bicycles & Invalid Carriages
Motor Vehicles		
Residential	Civil Engineering	
Non Residential	Special Trade Works	
Restaurants	Amusement and Recreational Services	Other Public Administration
Computer Services		
Water Transport	Banks	Other Financial Institution
Air Transport	Insurance	
Wholesale and Retail Trade	Land Transport	Professional
Financial Institution	Other Transport Services	Business Services
Public Administration	Port and Airport Operation Services	Education
Accommodation		Rental and Leasing
Health	Communication	Other Private Services
Defence and Public Order	Real Estate	

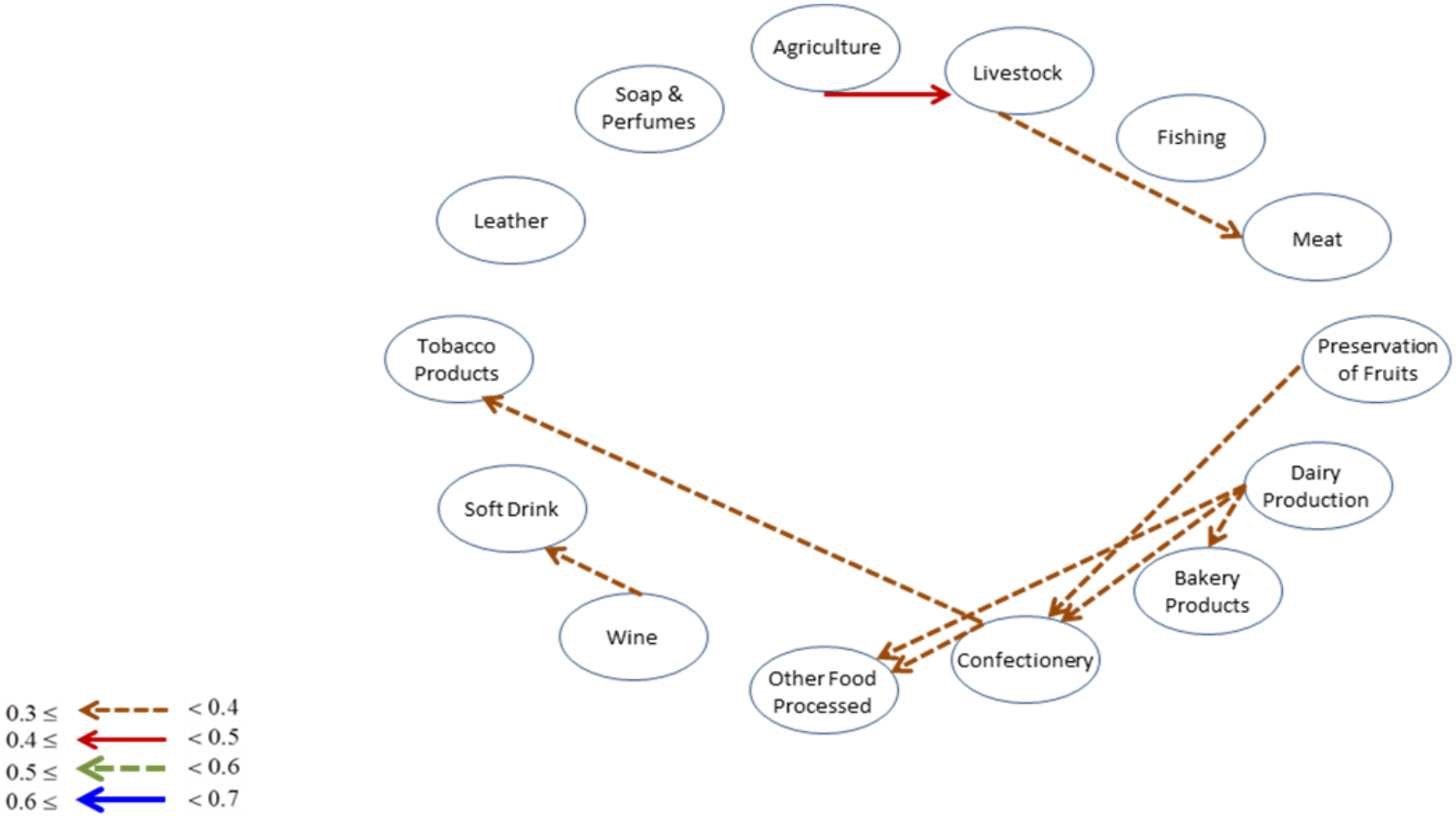
Findings

Functional Links among Industries in Cluster 1

- The relationship among the industries within a cluster is shown
- The structures of industrial clusters have strong correlation and inter-dependent relationship among industries. The directions of arrowheads denote the input-output relationship among industries, while the type of line denotes the strength of the supplier-buyer relationship. The correlation coefficient of industries in which the value exceeds 0.3 is displayed in the figures.

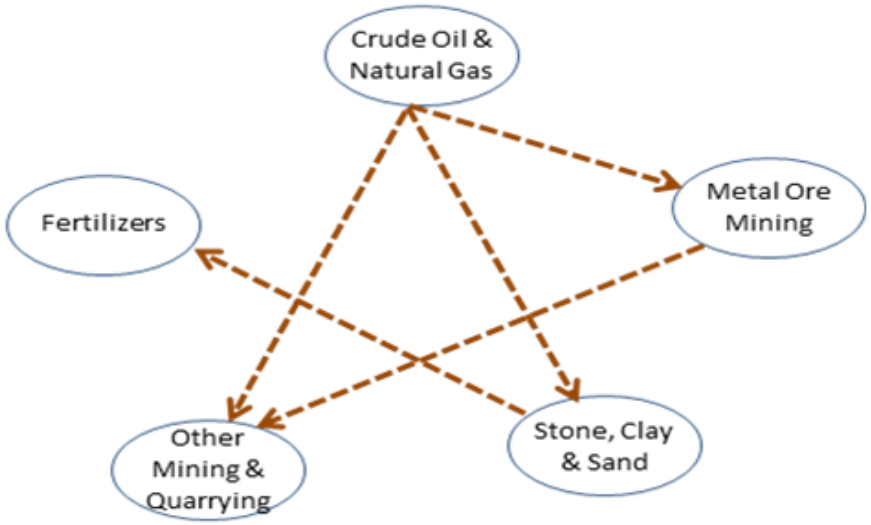
Findings

Functional Links among Industries in Cluster 2



Findings

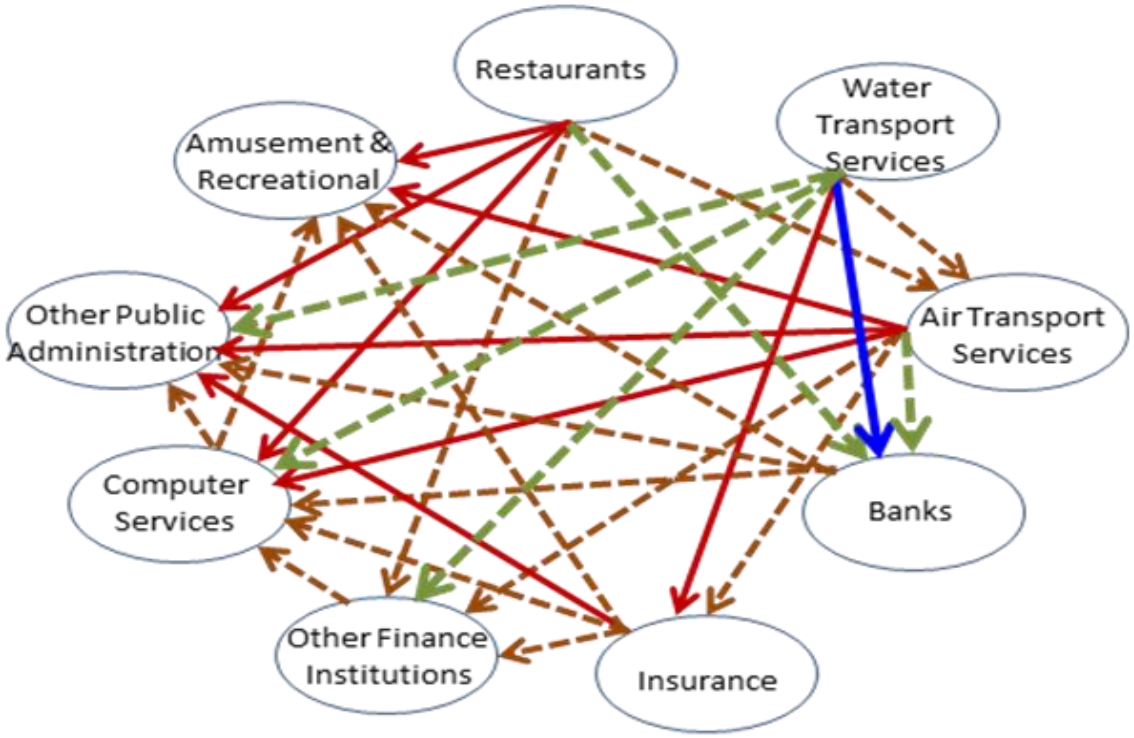
Functional Links among Industries in Cluster 3



0.3 ≤ < 0.4
0.4 ≤ < 0.5
0.5 ≤ < 0.6
0.6 ≤ < 0.7

Findings

Functional Links among Industries in Cluster 11



0.3 ≤ ⟵ < 0.4
0.4 ≤ ⟵ < 0.5
0.5 ≤ ⟵ < 0.6
0.6 ≤ ⟵ < 0.7

Conclusion

- Twelve (12) distinct industrial clusters were identified.
- Industries within a cluster have strong selling and purchasing relationships.
- This study is the first attempt to explore the similarities in selling and purchasing patterns of the industries in Malaysia using Input-Output table.
- Further research could include regional level data to identify industrial driver for each region.

Thank You