Unit 1 □ Point Pattern Analysis

Structure

1.1 Introduction

The flowchart of geographical understanding is : *Geographical Events about human habitat, economy and society → Spatial Pattern → Identification of Order in Observed Pattern → Determination of its Characteristics → Identification of the Order-forming Process → Multivariate Analysis → Scientific Geographical Explanation.* With roots in the 1950s, Quantitative Revolution (QR-I) actually took place the 1960s. The four main factors that promoted and maintained the development of a new approach in geography were – the availability of geographical data, the pace of change in the geographical phenomena, the technological changes in the handling of information and a pervasive belief in the usefulness of science. More emphasis was then given on the problem of why-man-lives-as-he-does and less emphasis on how he lives. This involves changes in methods by incorporating ideas from other disciplines. Geographers began talking of – spatial analysis, inferential techniques, concepts, laws, models, theories, behaviour, perception, prediction, ecosystem, linkages, matrices, equations, formulae and paradigms. Scientific explanation is given in-terms-of abstract mathematical and statistical parameters and therefore Quantitative Geographers viewed the human landscape in terms of set patterns, ordered processes and strict regularities.

Spatial differentiation is an important theme of geographical research. It concerns the study of differences between areas in terms of the numerous geographical phenomena (e.g., characteristics of slope, altitude, relief, soil, drainage, climate, vegetation, mineral, irrigation, agriculture/cropping, farming, industry, settlement and regional development and so on). In each case the ultimate aim is to find the amount of difference (i.e., how large is the difference) and explain it (i.e., how significant it is). Comparative statistics can be a great help in this regard : it provides a descriptive measure of the differences between sets of data and when the data relate to sample measurements, it also enable inferences to be made about differences between the populations from which the samples have been taken.

Requirements :

1. Database of Districtwise Population of West Bengal (1991)
2. Database of Districtwise Population of West Bengal (2001)
3. A Map of West Bengal with Districts
4. Ruler, Set Squares, Calculator, Transparent Graph (mm), Pen, Pencil, Eraser etc.

Procedures :

1. Draw a horizontal line touching the southern tip of the map, i.e., the x – axis
2. Draw a vertical line touching the western tip of the map, i.e., the y – axis
3. The intersection of the two produces the origin of the rectangular co-ordinate system
4. Graduate the lines in centimeter divisions
5. Locate, by eye estimation the centroid of the districts (i.e., the areal mean center of the districts)
6. Overlay a transparent graph paper on the map such that the rectangular reference system coincides with those on it
7. Derive by manipulation the rectangular co-ordinates, i.e., the x – and y – values (i.e., x_i, y_i) of the centroid of each district and record these neatly in a worksheet
8. Compute the products of (x_i) and (P_i) and (y_i) and (P_i) and then the co-ordinates of the mean centers of population separately for 1991 and 2001 as:

$$\bar{x}_p = [\Sigma(x_i.P_i)]/\Sigma P_i \text{ and}$$

$$\bar{y}_p = [\Sigma(y_i.P_i)]/\Sigma P_i$$

9. Locate the mean centres ($\bar{x}_p$, $\bar{y}_p$) on map by manipulation and label these appropriately
10. Measure using protractor and ruler the direction and magnitude of the shift of the mean centres during the 10 years

<u>**Note**</u> **:** Magnitude and direction of the locational shift of the mean centres can also be manipulated mathematically as :

$$\text{Magnitude (d)} = \sqrt{\{(\bar{x}_1 - \bar{x}_2)^2 + (\bar{y}_1 - \bar{y}_2)^2\}} \text{ and}$$

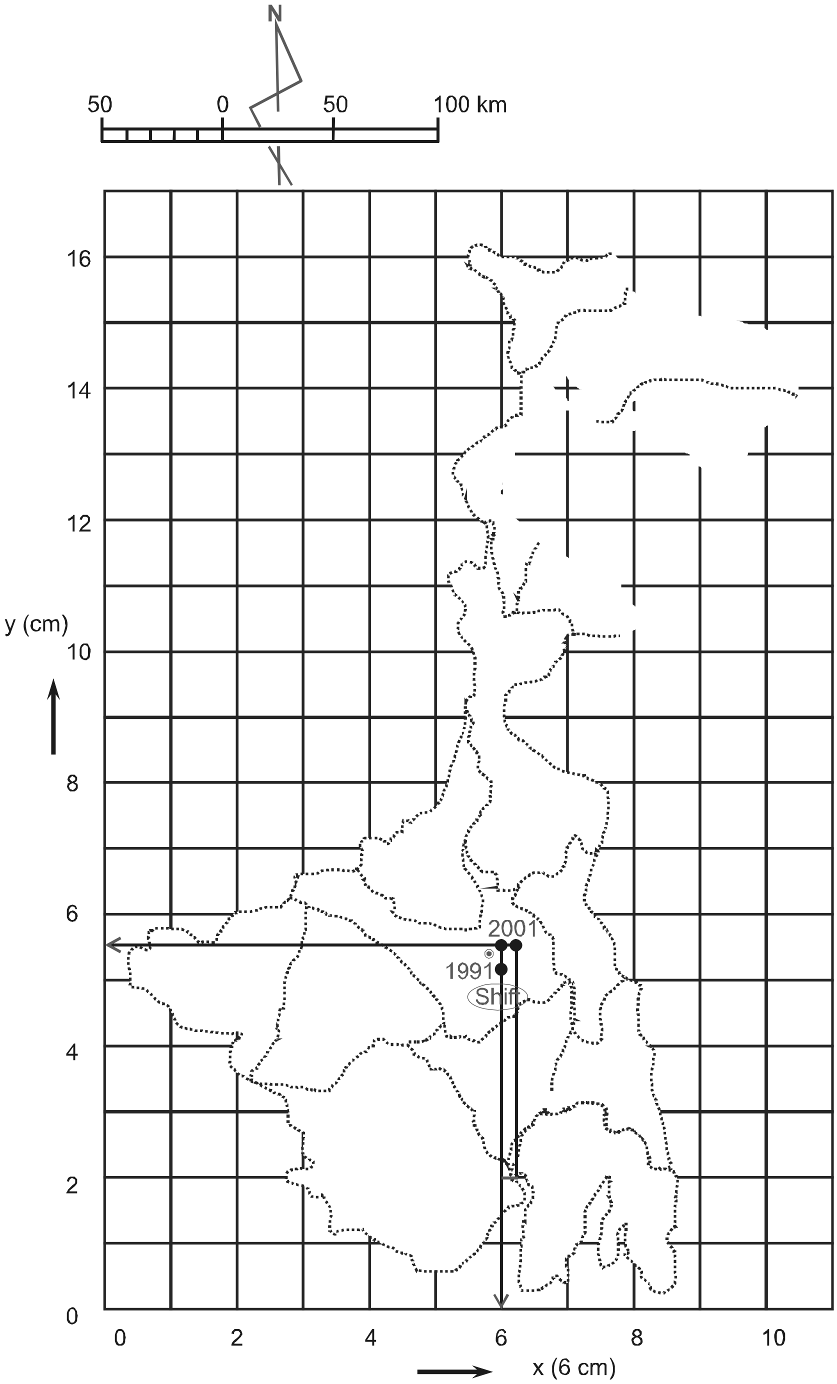
$$\text{Direction } (\theta) = \tan^{-1}\{(\bar{y}_1 - \bar{y}_2)/(\bar{x}_1 - \bar{x}_2)\}$$

11. Locations of the mean centers along with its shift are then interpreted with geographical perspective

Table – 1 : WORKSHEET FOR THE MEAN CENTRES OF POPULATION

District	Co-ordinates (cm)		Population		Products				Computations for Co-ordinates of Mean Centers
	x	y	2001 : (P_{01})	1991 : (P_{91})	$x.P_{01}$	$y.P_{01}$	$x.P_{91}$	$y.P_{91}$	
Purulia	1.6	5.0	2535233	2224577	4056372.8	12676165.0	3559323.2	11122885.0	
Bankura	3.5	4.5	3191822	2805065	11171377.0	14363199.0	9817727.5	12622792.5	
Midnapore	4.2	2.3	9638473	8331912	40481586.6	22168487.9	34994030.4	19163397.6	
Birbhum	4.8	6.7	3012546	2555664	14460220.8	20184058.2	12267187.2	17122948.8	
Burdwan	5.2	5.2	6919698	6050605	35982429.6	35982429.6	31463146.0	31463146.0	
Nadia	7.0	5.5	4603756	3852097	32226292.0	25320658.0	26964679.0	21186533.5	
U. Dinajpur	5.6	12.5	2441824	1926729	13674214.4	30522800.0	10789682.4	24084112.5	1991
Malda	6.0	10.0	3290160	2637032	19740960.0	32901600.0	15822192.0	26370320.0	$\overline{x}$ = (?(x.P_{91}) / ?P_{91})
Hooghly	5.8	3.8	5040047	4355230	29232272.6	19152178.6	25260334.0	16549874.0	= 5.97 cm
Howrah	5.8	2.8	4274010	3729644	24789258.0	11967228.0	21631935.2	10443003.2	$\overline{y}$ = (?(y.P_{91}) / ?P_{91})
Murshidabad	6.3	7.0	5863717	4740149	36941417.1	41046019.0	29862938.7	33181043.0	= 5.54 cm
Darjeeling	6.5	15.3	1605900	1299919	10438350.0	24570270.0	8449473.5	19888760.7	
Kolkata	6.7	3.0	4580544	4399819	30689644.8	13741632.0	29478787.3	13199457.0	2001
D. Dinajpur	7.0	10.8	1502647	1200924	10518529.0	16228587.6	8406468.0	12969979.2	
24 Parganas (S)	7.0	2.0	6909015	5715030	48363105.0	13818030.0	40005210.0	11430060.0	$\overline{x}$ = (?(x.P_{01}) / ?P_{01})
24 Parganas (N)	7.5	3.6	8930295	7281881	66977212.5	32149062.0	54614107.5	26214771.6	= 5.99 cm
Jalpaiguri	8.4	14.5	3403204	2800543	28586913.6	49346458.0	23524561.2	40607873.5	$\overline{y}$ = (?(y.P_{01}) / ?P_{01})
Coochbehar	9.0	13.5	2478280	2171145	22304520.0	33456780.0	19540305.0	29310457.5	= 5.60 cm
Sum Total =			80221171	68077965	480634675.8	449595642.9	406452088.1	376931415.6	

MEAN CENTRE OF POPULATION
WEST BENGAL, 2001



1.3 Nearest Neighbour Analysis of Settlement Pattern and its change over time

A geographical pattern is determined by the relative distances of spacings of a group of objects in relation to one another. Clark and Evans (1914) devised the Nearest Neighbour Analysis (NNA) based on the assumption that in a random pattern the first nearest neighbour distances are normally distributed. NNA concerns the measurement of actual straight-line distance irrespective of this distance with one that might be expected if the points were distributed at random in the same area.

The NN index is given by :

$$R_n = (\bar{r}_o/\bar{r}_e)$$

where,

$\bar{r}_o$ = observed mean nearest neighbour distance and $\bar{r}_e$ = expected mean nearest neighbour distance

The observed mean nearest neighbour distance, $\bar{r}_o = (\Sigma r_i)/n$

The expected mean nearest neighbour distance, $\bar{r}_e = (1/2). (\sqrt{}(A/n)$

where,

n = no. of settlements,

r_i = distance from i^{th} settlement to its nearest neighbour and

A = area

R_n provides and index of the departure from randomness. It is less, equal to or greater than one, depending upon whether the pattern tends to be aggregated, random, uniform & dispersed respectively.

Compute the Nearest Neighbour Index of Settlement Pattern and interpret.

Requirements :

1. A map showing spatial pattern of settlements
2. Ruler, Diagonal Scale, Divider, Protractor, Calculator, Pen, Pencil, Eraser etc.

Procedures :

1. Draw points at the centre of each settlement
2. Note the area of the map under consideration
3. Also note the total number of settlements
4. From each settlement measure the distance of its nearest neighbouring settlement

5. From the set of nearest neighbour distances, compute the mean observed nearest neighbour distance
6. Compute the mean expected nearest neighbour distance
7. Finally compute the nearest neighbour index and interpret the nature of spatial distribution of settlements

Table – 2 : WORKSHEET FOR NEAREST NEIGHBOUR ANALYSIS

<table>
<tr><th>Settlement with Nearest Neighbour</th><th>Nearest Neighbour Distance (cm)</th><th>Settlement with Nearest Neighbour</th><th>Nearest Neighbour Distance (cm)</th><th>Computation</th></tr>
<tr><td>1 – 2</td><td>1.25</td><td>23 - 27</td><td>2.60</td><td rowspan="21">?NND = 73.5 cm
Mean NND = $\bar{r}_o$
= (?NND / 44)
= 1.67 cm
= 0.835 Km

Density of Settlement:
P = (N / A)
= (44/72.285) / sq.km.
= 0.6087 / sq.km.

Mean Expected:
NND = $\bar{r}_e$
= 1 / (2* vP)
= 1 / (2* v0.608070)
= 0.641 Km

$R_n = (\bar{r}_o / \bar{r}_e)$
= (0.835 / 0.641)
= 1.30
Therefore,
the settlements are more random than dispersed.</td></tr>
<tr><td>2 – 1</td><td>1.25</td><td>24 - 25</td><td>1.60</td></tr>
<tr><td>3 – 2</td><td>2.10</td><td>25 - 24</td><td>1.60</td></tr>
<tr><td>4 – 5</td><td>2.00</td><td>26 - 21</td><td>1.40</td></tr>
<tr><td>5 – 6</td><td>1.70</td><td>27 - 31</td><td>1.70</td></tr>
<tr><td>6 – 5</td><td>1.70</td><td>28 - 25</td><td>1.80</td></tr>
<tr><td>7 – 8</td><td>2.70</td><td>29 - 30</td><td>1.20</td></tr>
<tr><td>8 – 7</td><td>2.70</td><td>30 - 35</td><td>1.00</td></tr>
<tr><td>9 – 11</td><td>2.20</td><td>31 - 36</td><td>1.20</td></tr>
<tr><td>10 - 14</td><td>1.00</td><td>32 - 33</td><td>1.10</td></tr>
<tr><td>11 - 17</td><td>1.70</td><td>33 - 32</td><td>1.10</td></tr>
<tr><td>12 - 16</td><td>0.90</td><td>34 - 29</td><td>1.30</td></tr>
<tr><td>13 - 16</td><td>2.50</td><td>35 - 30</td><td>1.00</td></tr>
<tr><td>14 - 15</td><td>1.00</td><td>36 - 31</td><td>1.20</td></tr>
<tr><td>15 - 14</td><td>1.00</td><td>37 - 31</td><td>1.20</td></tr>
<tr><td>16 - 12</td><td>0.90</td><td>38 - 35</td><td>1.70</td></tr>
<tr><td>17 - 11</td><td>1.70</td><td>39 - 40</td><td>1.60</td></tr>
<tr><td>18 - 19</td><td>1.40</td><td>40 - 39</td><td>1.60</td></tr>
<tr><td>19 - 18</td><td>1.40</td><td>41 - 42</td><td>3.30</td></tr>
<tr><td>20 - 44</td><td>2.10</td><td>42 - 41</td><td>3.30</td></tr>
<tr><td>21 - 26</td><td>1.40</td><td>43 - 38</td><td>2.20</td></tr>
<tr><td>22 - 27</td><td>2.50</td><td>44 - 15</td><td>1.70</td><td></td></tr>
</table>

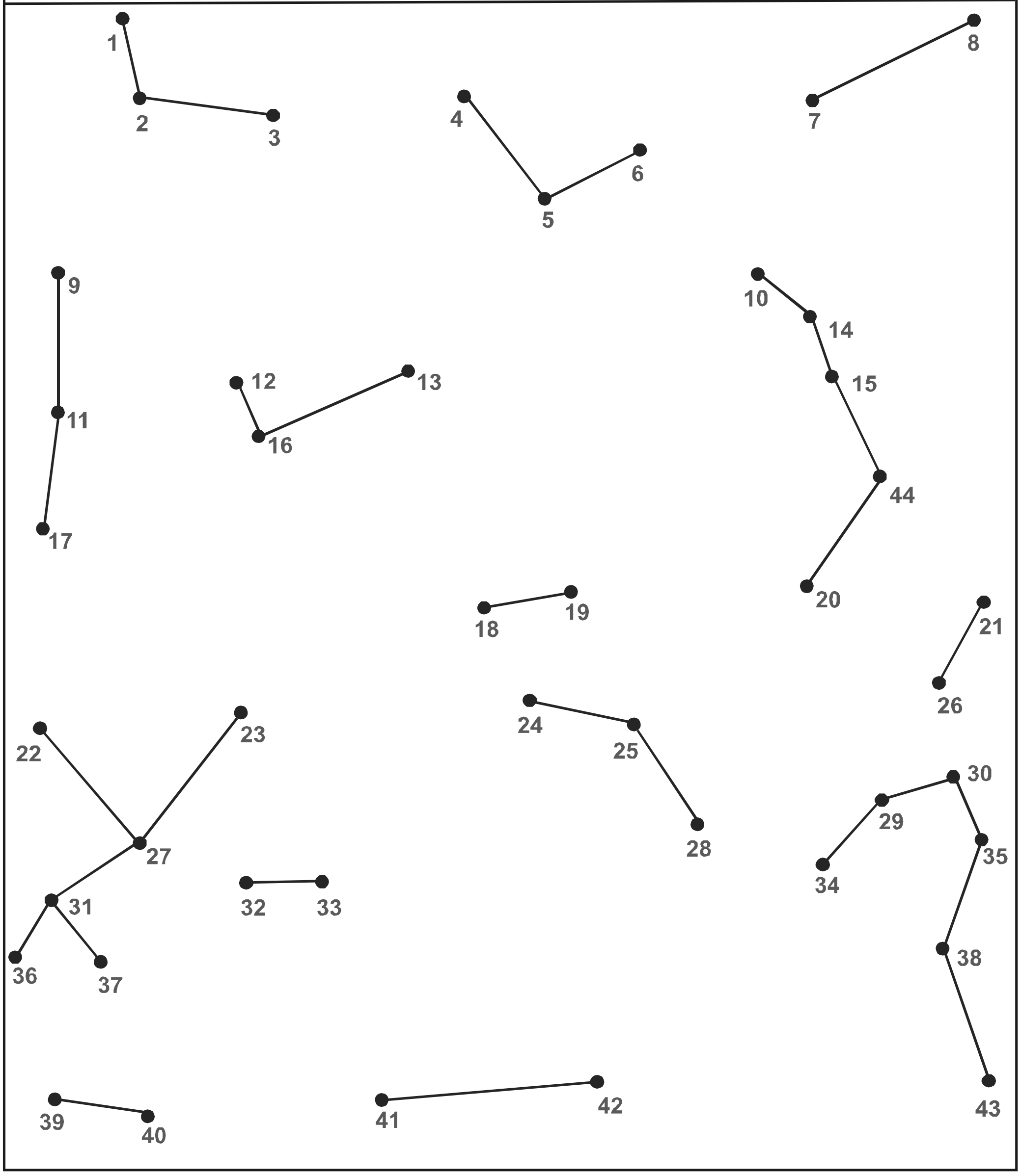
SETTLEMENT PATTERN ANALYSIS

BY

NEAREST NEIGHBOUR INDEX

TOPOGRAPHICAL MAP NO. - 73E / 2

GRID - B 3
SCALE = 1 : 50,000



Unit 2 ❑ Line Pattern Analysis

Structure

2.1 Measures of Connectivity (alpha, beta and gamma index)

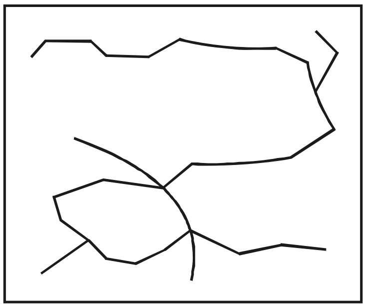
2.2 **Measures of accessibility from a point (de tour index)**

2.1 Measures of Connectivity (alpha, beta and gamma index)

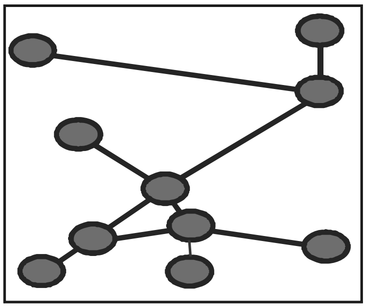
Several measure and indices can be used to analyze the network efficiency. Many of them were initially developed by Kansky, 1963 and can be used for

(a)expressing the relationship between values and the network structures they represent, (b) comparing different transportation networks at a specific point in time", and (c) comparing the evolution of a transport network at different points in time.

A graph is a symbolic representation of a network and of its connectivity. It implies an abstraction of the reality so it can be simplified as a set of linked nodes. In transport geography networks are analyzed for connectivity as all transport networks can be represented by graph theory in one way or the other. The following elements are fundamental at understanding graph theory :



Real Network

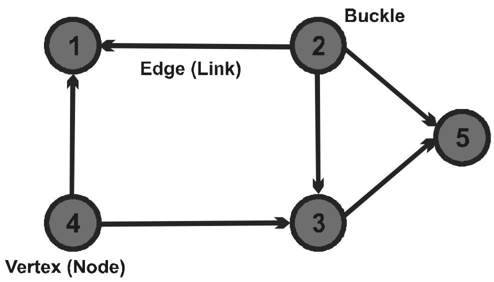


Graph Representation

Graph : A graph *G* is a set of vertex (nodes) v connected by edges (links) *e.* Thus *G=(v, e).*

Vertex (Node). A node v is a terminal point or an intersection point of a graph. It is the abstraction of a location such as a city, an administrative division, a road intersection or a transport terminal (stations, terminuses, harbors and airports).

Edge (Link). An edge *e is a* link between two nodes. The link *(i, j)* is of initial extremity *i* and of terminal extremity *j. A* link is the abstraction of a transport infrastructure supporting movements between nodes. It has a direction that is commonly represented as an arrow. When an arrow is not used, it is assumed the link is bidirectional.

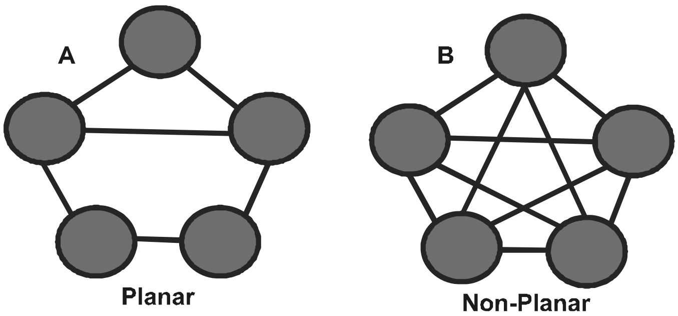


Sub-Graph. A sub-graph is a subset of a graph *G* where *p* is the number of subgraphs. For instance $G' = (v,e')$ can be a distinct sub-graph of *G*. Unless the global transport system is considered in its whole, every transport network is in theory a sub-graph of another. For instance, the road transportation network of a city is a subgraph of a regional transportation network, which is itself a sub-graph of a national transportation network.

Buckle. A link that makes a node correspond to itself is a buckle.

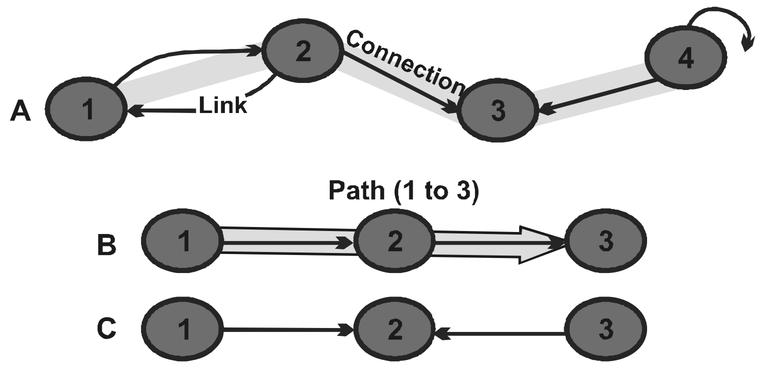
Planar Graph. A graph where all the intersections of two edges are a vertex. Since this graph is located within a plane, its topology is two-dimensional.

Non-planar Graph. A graph where there are no vertex at the intersection of at least two edges. This implies a third dimension in the topology of the graph since there is the possibility of having a movement "passing over" another movement such as for air transport. A non-planar graph has potentially much more links than a planar graph.



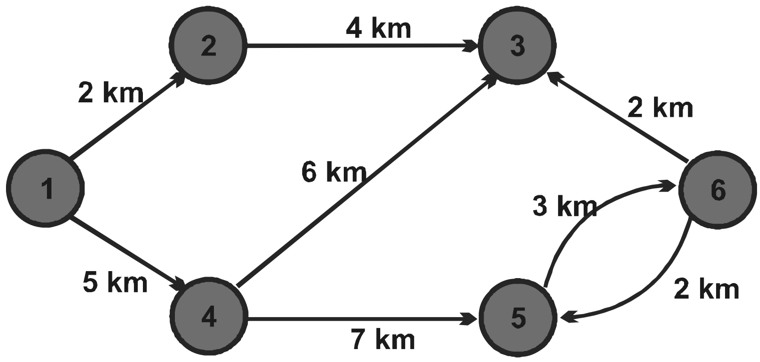
Graph A is planar since no linkage is overlapping with another. Graph B is non-planar since many links are overlapping and cannot be reconfigured in a manner that would make it planar. The goal of a graph is representing the structure, not the appearance of a network. The conversion of a real network into a planar graph is a straightforward process which follows some basic rules : (a) every terminal and intersection point becomes a node, (b) each connected nodes is then linked by a straight segment. The real network, depending on its complexity, may be confusing in terms of revealing its connectivity (what is linked with what). A graph representation reveals the connectivity of a network in the best possible way. Other rules can also be applied, depending on the circumstances :

*A node that is not a terminal or an intersection point can be added to the graph if along that segment an **attribute is changing.** For instance, it would be recommended to represent as a node the shift from 2 lanes to 4 lanes along a continuous road segment, even if that shift does not occur at an intersection or terminal point.*



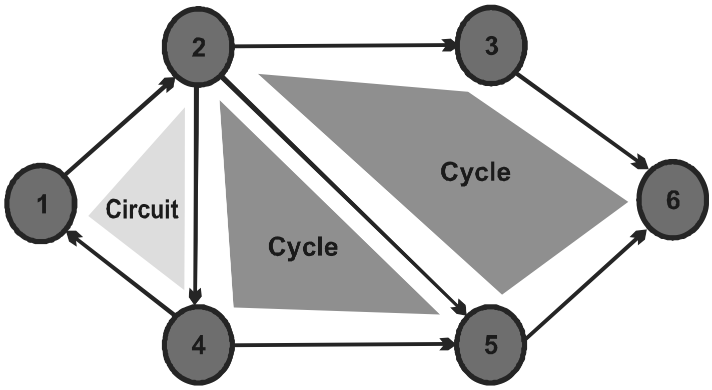
Connections and Paths

On graph A, there are 5 links [(1,2), (2,1), (2,3), (4,3), (4,4)] and 3 connections [(1-2), (2-3), (3-4)]. On graph C, there is a path between 1 and 3, but on graph C there is no path between 1 and 3.



Length of a Link, Connection or, Path

On this graph, the length of link (2,3) is 4 km awl the length of the path between 1 and 6 (1-4-5-6; 3 segments) is 15 km.

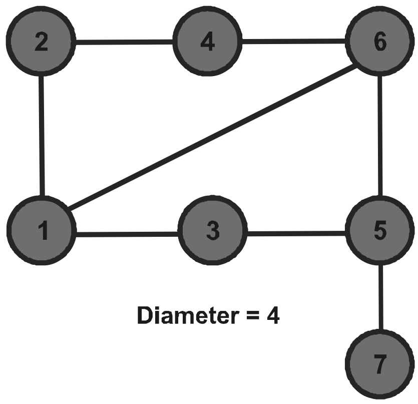


Cycles and Circuits

On this graph, 2-3-6-5-2 is a cycle but not a circuit. 1-2-4-1 is a cycle and a circuit.

The three basic measures of the structural attributes of a graph are the diameter, the number of cycles and the order of a node.

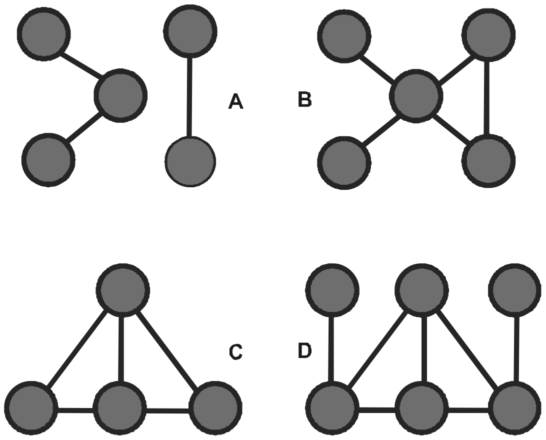
Diameter *(d)* : It is defined as the length of the shortest path between the most distanced nodes of a graph. Thus, it measures the extent of a graph and the topological length between two nodes. It also enables to measure the development of a network in time. The larger the diameter, the less linked a network tends to be.



Shibel Distance							
v	1	2	3	4	5	6	7
1	0	1	1	2	2	1	3
2	1	0	2	1	3	2	4
3	1	2	0	3	1	2	2
4	2	1	3	0	2	1	3
5	2	3	1	2	0	1	1
6	1	2	2	1	1	0	2
7	3	4	2	3	1	2	0

Number of Cycles ***(u)*** : It is defined as the maximum number of independent cycles in a graph and is estimated through the number of nodes (v), links *(e)* and of sub-graphs (p). The formula is, $u = e\text{-}v+p$.

Thus, in simple networks, $u = 0$ since they have no cycles. The more complex a network is, the higher the value of *u,* so it can be used as an indicator of the level of development and complexity of a transport system.



$u = e - v + p$

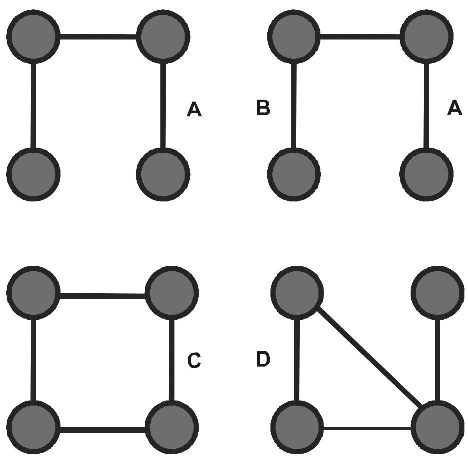
	e	v	p	u
A	3	5	2	0
B	5	5	1	1
C	5	4	1	2
D	6	7	1	2

<u>Order</u> <u>(degree) of a Node</u> (*o*) : It is defined as the number of its attached links and is a simple, but effective measure of nodal importance. The higher its value, the more a node is important in a graph as many links converge to it. Hub nodes have a high order, while terminal points have an order that can be as low as 1. A perfect hub would have its order equal to the summation of all the orders of the other nodes in the graph and a perfect spoke would have an order of 1.

Connectivity Indexes : Indexes are more complex methods to represent the structural properties of a graph since they involve the comparison of a measure over another :

<u>Beta Index</u> (β) : It measure the level of connectivity in a graph and is expressed *by* the relationship between the number of links (e) over the number of nodes (v). For trees and simple networks, β < **1** and for a connected network with me cycle, β = **1**. More complex networks have a value greater than 1. In a network with a fixed number of nodes, the higher the number of links, the higher the number of paths possible in the network. Therefore, complex networks have a high value of Beta.

$\beta = \frac{e}{v}$

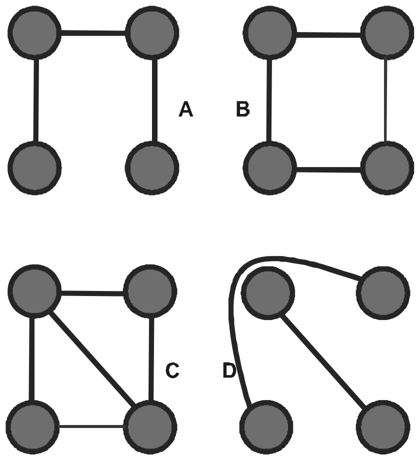


$\beta = \frac{e}{v}$

	e	v	Beta
A	2	4	0.5
B	3	4	0.75
C	4	4	1.0
D	5	4	1.25

Alpha Index (a) : It is a measure of connectivity which evaluates the number of cycles in a graph in comparison with the maximum number of cycles. The higher the alpha index, the more a network is connected. For trees and simple networks, $a = 0$ and for a completely connected network, a = 1. Thus, it measures the level of connectivity independently of the number of nodes. It is very rare that a network will have an alpha value of 1, because this would imply very serious redundancies.

$$a = \frac{u}{2v - 5}$$

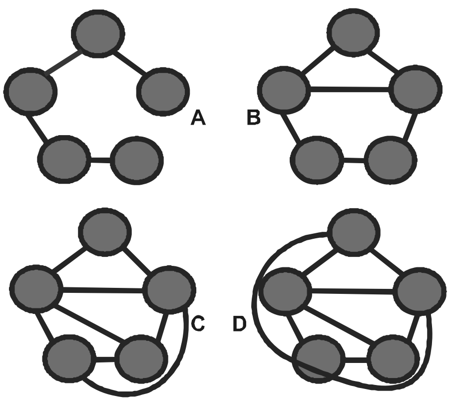


$$a = \frac{u}{2v - 5}$$

	u(e-v+p)	2v-5	Alpha
A	0	3	0.0
B	1	3	0.33
C	2	3	0.66
D	3	3	1.0

Gamma Index (y): It is a measure of connectivity that considers the relationship between the number of observed links and the number of possible links. The value of gamma lies between 0 and 1. $\gamma = 1$ indicates a completely connected network, although it is extremely unlikely in reality. Gamma is an efficient value to measure the progression of a network in time.

$$\gamma = \frac{\mathbf{e}}{\mathbf{3(v - 2)}}$$



$$\gamma = \frac{3}{3(v - 2)}$$

	e	3(v-2)	Gamma
A	4	9	0.44
B	6	9	0.66
C	8	9	0.88
D	9	9	1.0

Example : Compute the attributes of the structure of the given transport network and interpret.

Requirements :

1. A Transport Network Map
2. Calculator, Marker, Pen, Pencil, Eraser etc.

Procedures :

1. Identify and count the numbers of vertex, edge, subgraphs, planar and non-planar graphs
2. Compute the diameter, number of cycles and order of nodes
3. Compute the alpha, beta and gamma index and interpret the values

Table – 3 : WORKSHEET FOR CONNECTIVITY INDICES

Transport Network of Roadways of South Bengal

Total No. of 'arc' = 95
Total No. of 'node' = 71

alpha index	**= (arc – node – 1) / (2 . node – 5)**
	= (95 – 71 – 1) / (2 x 71 – 5)
	= 23 / 137
	= 0.17

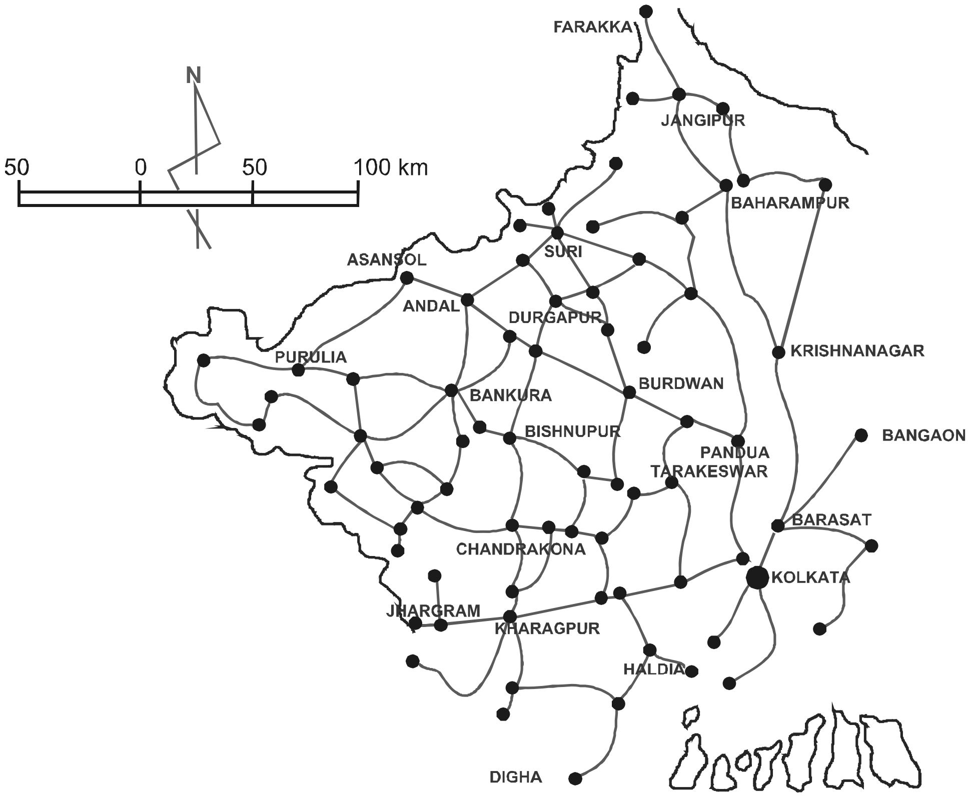
beta index	**= arc / node**
	= 95 / 71
	= 1.34

gamma index	**= arc / 3. (node – 2)**
	= 95 / 3 x (71 – 2)
	= 95 / 207
	= 0.46

TRANSPORT NETWORK ANALYSIS

OF

THE ROADWAY NETWORK OF SOUTH BENGAL



DEGREE OF CONNECTIVITY

arc=95, node=71

μ = arc- (node-1)
= 95 - (71-1)
= **25**

α = (arc-node-1)/(2. node-5)
= 95-71-1)/[(2).(71)-5]
= **0.17**

β = arc/node
= 95/71
=**1.34**

γ = arc/3 (node-2)
= 95/3 (71-2)
= **0.46**

2.2 Measures of accessibility from a point (de Tour index)

Detour Index : It is a measure of the efficiency of a transport network in terms of how well it overcomes distance or the friction of distance. The closer the detour index gets to 1, the more the network is spatially efficient. Networks having a detour index of 1 are rarely, if ever, seen and most networks would fit on an asymptotic curve getting close to 1, but never reaching it.

$$DI = \frac{DD}{TD}$$

For instance, the straight distance (*DD*) between two nodes may be 40 km but the transport distance (*TD*; real distance) is 50 km. The detour index is thus 0.8 (40/50). The complexity of the topography is often a good indicator of the level of detour.

Example : Draw a 'de tour map' for the given Road Transport Network and interpret.

Requirements :

1. A Map showing Network of Road Transport
2. Calculator, Marker, Pen, Pencil, Eraser etc.

Procedures :

1. From the central point find the straight distance (DD)
2. From the central point find also the real transport distance (TD)
3. Compute the de Tour index for all the given nodes, except the central node and jot down the values for respective locations
4. Draw isopleths of de Tour index with a suitable interval and interpret.

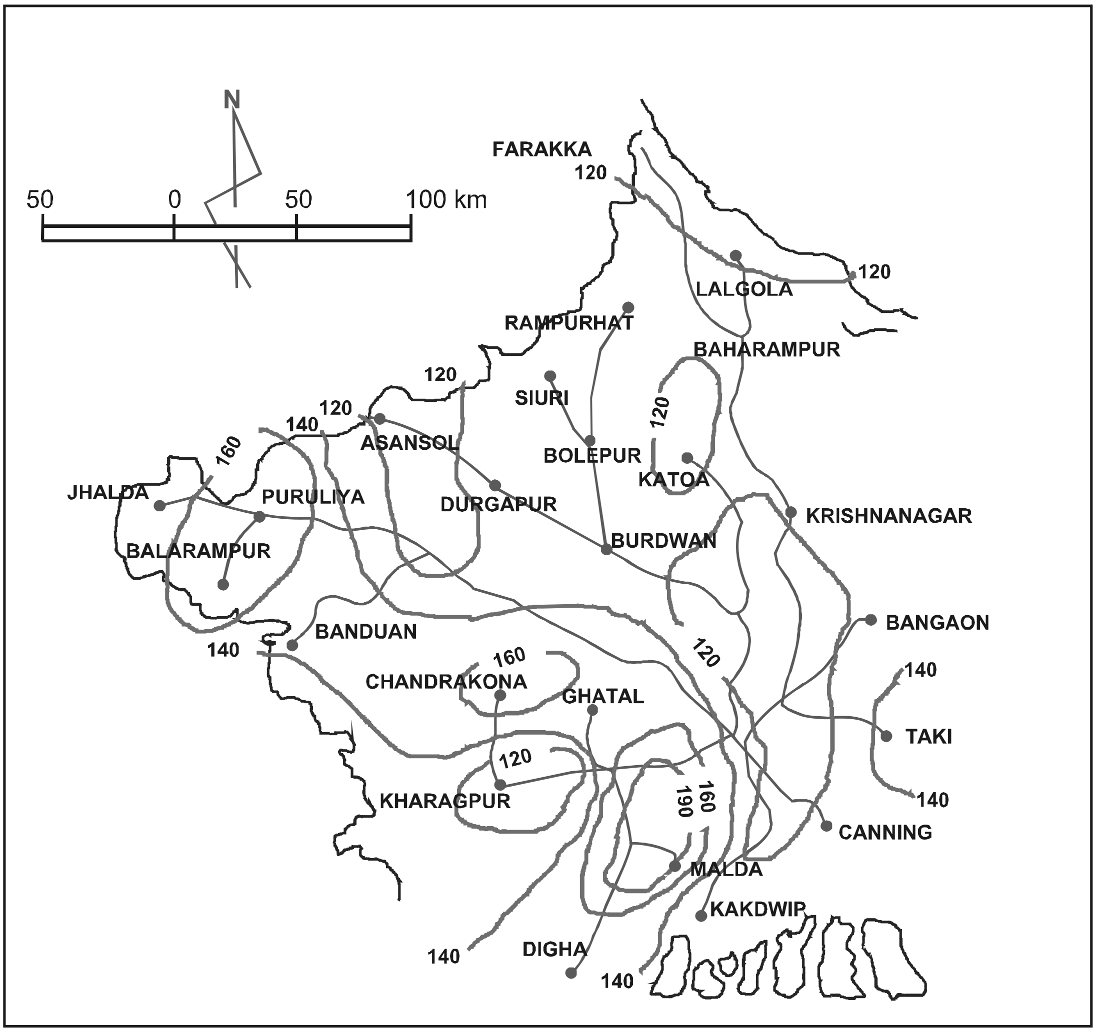
Table – 4 : WORKSHEET FOR DE TOUR INDEX

Base Point : Kolkata

Place	**Actual Distance (km) : TD**	**Direct Distance (km) : DD**	**De Tour Index = (TD / DD) x 100 %**
Canning	047.2	038.9	121.3
Taki	069.4	047.2	147.0
Bongaon	088.8	072.2	123.0
Krishnanagar	113.0	091.7	123.2
Baharampur	202.0	163.9	123.2
Lalgola	228.0	197.2	115.6
Farakka	288.8	241.7	119.5
Rampurhat	227.8	180.6	126.1
Suri	205.0	163.9	125.0
Bolpur	163.0	133.3	122.2
Burdwan	119.0	091.7	129.7
Durgapur	182.0	144.4	126.0
Asansol	222.0	191.7	115.8
Bankura	175.0	147.2	118.9
Purulia	328.0	197.2	166.3
Jhalda	372.0	247.2	150.5
Balarampur	358.0	213.9	167.3
Banduan	258.3	183.3	140.9
Chandrakona	166.6	100.0	166.6
Kharagpur	118.0	102.8	114.7
Digha	170.0	113.9	149.2
Haldia	113.8	052.8	215.5
Ghatal	100.0	063.9	156.4
Katwa	138.8	116.6	119.0
Kakdwip	077.8	063.8	121.9

TRANSPORT NETWORK ANALYSIS
ACCESSIBILITY MAP BASED ON DE TOUR INDEX

(BASE POINT = KOLKATA)
(ROADWAY NETWORK OF SOUTH BENGAL)



Unit 3 □ Areal Pattern Analysis

3.1 Measures of specialisation

3.1.1 Dominant and Distinctive Analysis

This is an important technique used in geographical studies to identify within a group of attributes the dominant one as well as the distinctive one. The dominant one of a geographical object or entity can be easily identified from the highest percentage share of an attribute. The distinctive one is judged based on a set of attribute values for the whole region. For this, 'mean' and 'standard deviation' are used to determine the 'degree of distinctiveness' of any particular attribute. For this, a worksheet to be prepared with geographical entities (i.e., blocks, districts. States, etc) in the 1st column and the values of the attributes in the successive columns. The values may be absolute numbers as well as relative proportions (%).

The following table shows a distribution of proportion of three (3) variables in the six (6) districts of a state. Figures corresponding to rows concern a particular district and are used to identify the dominant attribute of a district. Thus attribute – 1 is dominant in AA, BB, and DD districts, attribute – 2 is dominant in CC and FF districts and attribute – 3 is dominant only in EE district. Thus dominant one is identified by inspecting the rowwise values of a geographical entity.

District	Values on Attributes (%)		
	Variable – 1	Variable – 2	Variable – 3
AA	56	30	14
BB	68	20	12
CC	32	42	26
DD	40	30	30
EE	10	20	70
FF	30	60	10
Mean	39.3	33.7	27.0
Standard Deviation	20.5	15.3	22.5

Attribute	Criteria	Range	Degree of Distinctiveness	District
1	(Mean + 1SD) to (Mean + 2SD)	59.8 – 80.3	1	BB
2	(Mean + 1SD) to (Mean + 2SD)	49.0 – 64.3	1	FF
3	(Mean + 1SD) to (Mean + 2SD)	49.5 – 72.0	1	EE
Thus, attribute – 1 is distinctive only in BB district, attribute – 2 in FF district and attribute – 3 in EE district.				

The distinctiveness of an attribute is identified by analysing the values entered in a particular column. This means distinctiveness concerns a single attribute as distributed within the six districts of a state. For this 'mean' and 'standard deviation' of the attributes are computed first and then compared within the districts. Degree of distinctiveness is 1, if the attribute value lies between (Mean + 1SD) to (Mean + 2SD), 2 when the value lies between (Mean + 2SD) to (Mean + 3SD), 3 when (Mean + 3SD) to (Mean + 4SD), and so on. Thus dominant is **'local'** but distinctiveness is **'regional'** in character and context. **All distinctive attributes may be dominant but all dominants are never distinctive.**

Example : Make a 'Dominant and Distinctive Analysis' of Crop Production of West Bengal, 2001 - 2002.

Requirements :

1. A database with districtwise data of crop production.
2. Calculator, Marker, Pen, Pencil, Eraser etc.

Procedures :

1. Find the highest value in each row, mark them as the 'dominant' one
2. For each column, compute the 'mean' and 'standard deviation'
3. Prepare a 'distinctiveness table' and for each landuse category identify the 'distinctive' ones with 'degree'
4. Draw 'diagrammatic map' to show the proportional distribution of crop production and mark the 'degree of distinctiveness' with appropriate symbols (Category = letter symbols; Degree = numerals)

Table – 5 : WORKSHEET FOR DOMINANT AND DISTINCTIVE ANALYSIS

Districts	Production of Crops ('000 tonnes)					Proportion of Production of Crops (% of Total)				
	Rice	Pulses	Oilseeds	Jute	Potato	Rice	Pulses	Oilseeds	Jute	Potato
24 Parganas (N) : J	856.4	4.8	37.7	886.6	138.1	**44.5**	0.2	2.0	**46.1**	7.2
24 Parganas (S): R	1003.7	6.6	4.6	16.6	83.5	**90.0**	0.6	0.4	1.5	7.5
Bankura	1222.4	0.4	18	2.4	522.9	**69.2**	0.0	1.0	0.1	29.6
Birbhum : R, P	1157.4	17.9	38.8	7.7	273.2	**77.4**	**1.2**	2.6	0.5	18.3
Burdwan	1930.6	3.5	44.5	400.9	1217.3	**53.7**	0.1	1.2	11.1	33.8
Coochbehar : J	466.7	5.7	7.2	904.3	275.1	28.1	0.3	0.4	**54.5**	16.6
D. Dinajpur	425.3	2.2	17.1	182.5	66.5	**61.3**	0.3	2.5	26.3	9.6
Darjeeling : Po	59.2	1.3	0.3	30.9	98.6	31.1	0.7	0.2	16.2	**51.8**
Hooghly : Po_3	847.7	0.3	27.6	623.5	2356.4	22.0	0.0	0.7	16.2	**61.1**
Howrah	287.6	0.2	3.7	80.9	198.4	**50.4**	0.0	0.6	14.2	34.8
Jalpaiguri	410.1	1.9	6.3	510.3	314.9	33.0	0.2	0.5	**41.0**	25.3
Malda : P_3	536.2	25.6	37	370.8	44.5	**52.9**	**2.5**	3.6	36.6	4.4
Midnapore	2710.8	14.4	79.5	167.1	1823.4	**56.5**	0.3	1.7	3.5	38.0
Murshidabad : J	1085.4	34.7	58.6	2010	174.4	32.3	1.0	1.7	**59.8**	5.2
Nadia : J, O_2	958.7	38.4	1116.4	2330.6	106.2	21.1	0.8	**24.5**	**51.2**	2.3
Purulia : R_2	745.1	6.5	2.3	0	12.7	**97.2**	0.8	0.3	0.0	1.7
U. Dinajpur : Po	554.4	10.7	29.7	611.1	1161.3	23.4	0.5	1.3	25.8	**49.1**
Note : Dominant Crops = underlined Bold. Distinctive Crops = in colour (Red = 1st order, Blue = 2nd order, Pink = 3rd order)										
					Mean	49.7	0.6	2.7	23.8	23.3
					Stand. Dev	23.6	0.6	5.7	20.8	18.9

Dominant Crops :

(1) Rice – North 24 Parganas, South 24 Parganas, Bankura, Birbhum, Burdwan, D. Dinajpur, Howrah, Malda, Midnapur and Purulia

(2) Jute – Coochbehar, Jalpaiguri, Murshidabad and Nadia

(3) Potato – Darjilling, Hooghly and U. Dinajpur

Distinctive Crops :

(1) Rice - South 24 Parganas and Birbhum (1st order); and Purulia (2nd order)

(2) Pulses – Birbhum (1st order); and Malda (3rd order)

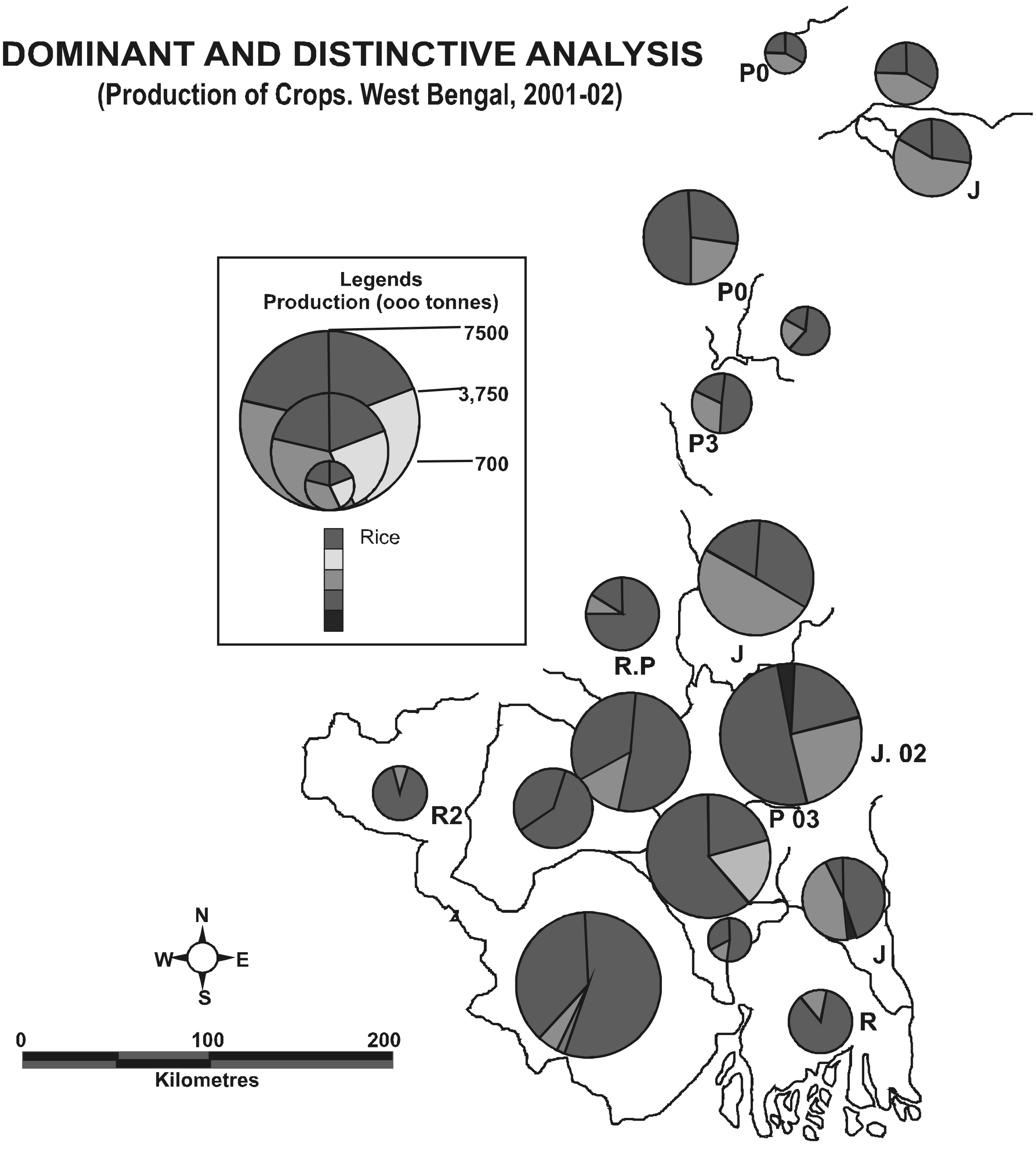
(3) Oilseeds – Nadia (3rd order)

(4) Jute – North 24 Parganas, Coochbehar, Murshidabad and Nadia (1st order)

(5) Potato – Darjilling, U. Dinajpur (1st order); and Hooghly (3rd order)

DOMINANT AND DISTINCTIVE ANALYSIS

(Production of Crops. West Bengal, 2001-02)



3.1.2 : Indices of Specialization Location Quotient

The **location quotient technique** is most commonly utilized economic base analysis method **(Haig, 1928)**. It compares the local economy to a reference economy and in the process attempts to identify specializations in the local economy. Location quotients are calculated for all industries to determine whether or not the local economy has a greater share of each industry than expected when compared to a reference economy. If an industry has a greater share than expected of a given industry, then that "extra" industry employment is assumed to be basic because those jobs are above what a local economy should have to serve local needs. The location quotient is most frequently used in economic geography and locational analysis, but it has much wider applicability. The location quotient (LQ) is an index for comparing an area's share of a particular activity with the area's share of some basic or aggregate phenomenon. The formula for computing location quotients can be written as :

$$LQ = \frac{e_i / e}{E_i / E}$$

where,

e_i = Local employment in industry in year T

e = Total local employment in year T

E_i = National employment in industry in year T

E = Total national employment in year T

In this formula, the 'regional' or 'local' economy (often a county or district or state) to the 'national' economy. Location quotients may also be calculated that compare the county to a state. The LQ provides evidence for the existence of basic employment in a given industry. Interpreting the Location Quotient is very simple. Only three general outcomes are possible when calculating location quotients. These outcomes are as follows :

LQ < 1.0 means all employment is non-basic, i.e., local employment is less than was expected for a given industry. In other words, the area has less of a share of the activity than is more generally, or regionally, found. Therefore, that industry is not even meeting local demand for a given good or service. Therefore all of this employment is considered non-basic by definition.

LQ = 1.0 means all employment is non-basic, i.e., local employment is exactly sufficient to meet the local demand for a given good or service. In other words, the area has a share of the activity in accordance with its share of the base. Therefore, all of this employment is also considered non-basic because none of these goods or services are exported to non-local areas.

LQ > 1.0 means some employment is basic, i.e., local employment is greater than expected and it is therefore assumed that this "extra" employment is basic. In other words, there is a relative concentration of the activity in the area compared to the region as a whole. These extra jobs then must export their goods and services to non-local areas which, by definition, makes them basic sector employment.

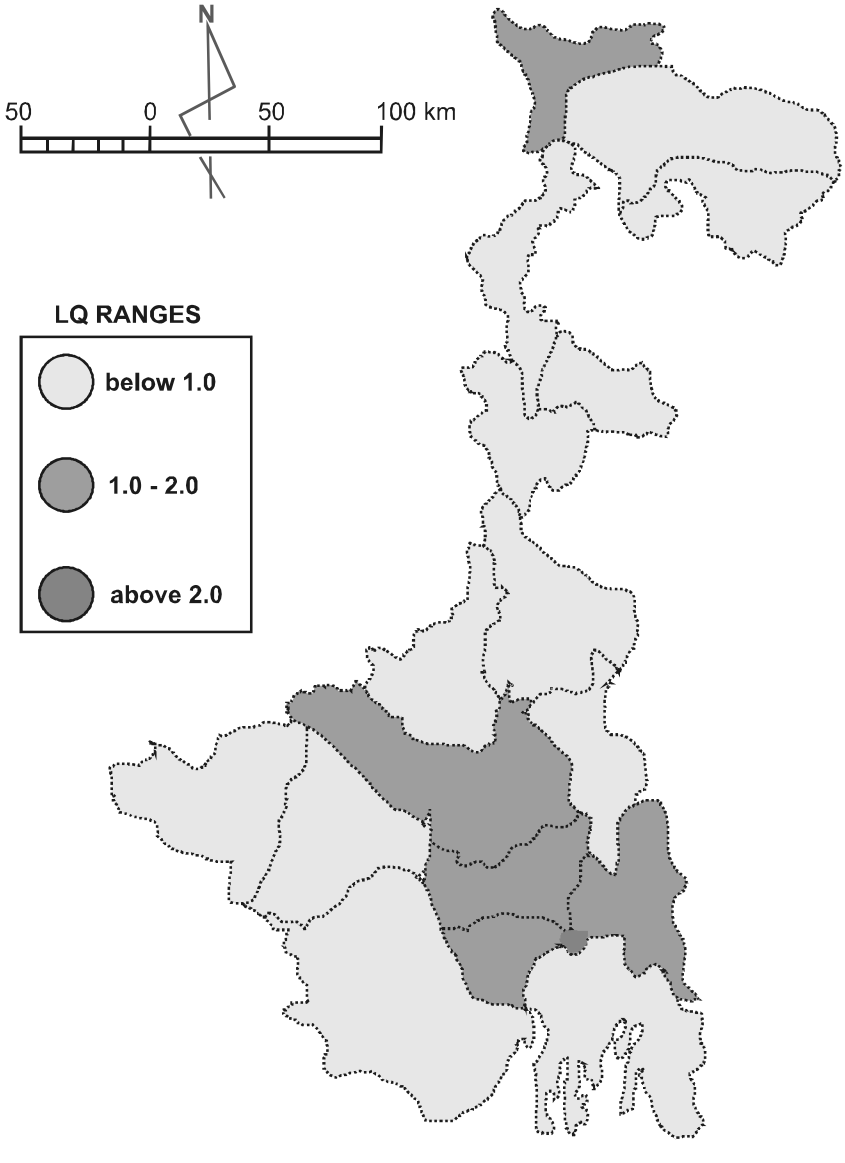
Table – 6 : WORKSHEET FOR LOCATION QUOTIENTS

Districts	Urban Population, 2001 : U	Population 2001 : P	U / TU	P / TP	LQ= (U / TU) / (P / TP)
Purulia	255239	2535233	0.01135	0.03160	0.36
Bankura	235264	3191822	0.01046	0.03978	0.26
Midnapore	1010954	9638473	0.04495	0.12014	0.37
Birbhum	258479	3012546	0.01149	0.03755	0.31
Burdwan	2572423	6919698	0.11439	0.08625	1.33
Nadia	979047	4603756	0.04353	0.05738	0.75
U. Dinajpur	294471	2441824	0.01309	0.03043	0.43
Malda	240915	3290160	0.01071	0.04101	0.26
Hooghly	1687410	5040047	0.07504	0.06282	1.19
Howrah	2153571	4274010	0.09575	0.05327	1.79
Murshidabad	732343	5863717	0.03256	0.07309	0.44
Darjeeling	520877	1605900	0.02316	0.02001	1.16
Kolkata	4580544	4580544	0.20307	0.05709	3.56
D. Dinajpur	196643	1502647	0.00874	0.01873	0.47
24 Parganas (S)	1089730	6909015	0.04846	0.08612	0.56
24 Parganas (N)	4849218	8930295	0.21565	0.11132	1.94
Jalpaiguri	603847	3403204	0.02685	0.04242	0.63
Coochbehar	225506	2478280	0.01002	0.03089	0.32
	TU=22486481	TP=80221171			

Table – 6a : DISTRIBUTION OF LQs

Range of LQs	Districts	Remarks
below 1	Purulia, Bankura, Midnapur, Birbhum, Nadia, Malda, U. Dinajpur, D. Dinajpur, Murshidabad, South 24 Parganas, Jalpaiguri, Cooch Behar	Low
1 to 2	Burdwan, Hooghly, Howrah, Darjilling, North 24 Parganas	Moderate
Above 2	Kolkata	High

CONCENTRATION
OF
URBAN POPULATION OF WEST BENGAL, 2001
BY
LOCATION QUOTIENT ANALYSIS



F - 15

Example : Compute the Location Quotients of Urban Population of West Bengal, 2001 and interpret the Location Quotient Map.

Requirements :

1. A database with districtwise Urban and Total Population of West Bengal
2. Calculator, Marker, Pen, Pencil, Eraser etc.

Procedures :

1. Prepare a work sheet with the following columns : Districts, Urban Population of a district (U), Total Population of a district (T), (U/T_U), (P/T_P), and LQ $= (U/T_U)/(P/T_P)$, where Total Urban Population of West Bengal = T_U and Total Population of West Bengal $=T_P$
2. Compute and enter the figures for 4th, 5th and 6th columns

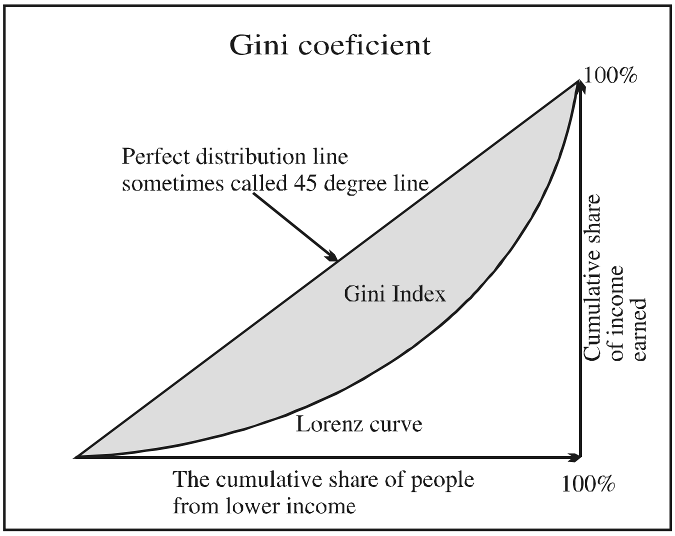
Districts	Urban Population 2001 (U)	Population 2001(P)	(U/T_U)	(P/T_P)	LQ $= (U/T_U)/(P/T_P)$
Purulia	255239	2535233	0.01135	0.03160	0.36
Bankura	235264	3191822	0.01046	0.03978	0.26
Midnapore	1010954	9638473	0.04495	0.12014	0.37
Birbhum	258479	3012546	0.01149	0.03755	0.31
Burdwan	2572423	6919698	0.11439	0.08625	1.33
Nadia	979047	4603756	0.04353	0.05738	0.75
U. Dinajpur	294471	2441824	0.01309	0.03043	0.43
Malda	240915	3290160	0.01071	0.04101	0.26
Hooghly	1687410	5040047	0.07504	0.06282	1.19
Howrah	2153571	4274010	0.09575	0.05327	1.79
Murshidabad	732343	5863717	0.03256	0.07309	0.44
Darjeeling	520877	1605900	0.02316	0.02001	1.16
Kolkata	4580544	4580544	0.20307	0.05709	3.56
D. Dinajpur	196643	1502647	0.00874	0.01873	0.47
24 Parganas (S)	1089730	6909015	0.04846	0.08612	0.56
24 Parganas (N)	4849218	8930295	0.21565	0.11132	1.94
Jalpaiguri	603847	3403204	0.02685	0.04242	0.63
Coochbehar	225506	2478280	0.01002	0.03089	0.32
	$\Sigma = 22486481$	$\Sigma = 80221171$			

Range of LQ	Districts	Remarks
Below 1	Purulia, Bankura, Midnapore, Birbhum, Nadia, Malda, U. Dinajpur, D. Dinajpur, Murshidabad, 24 Parganas (S), Jalpaiguri, Coochbehar	Low
1 – 2	Burdwan, Hooghly, Howrah, Darjeeling, 24 Parganas (N)	Moderate
Above 2	Kolkata	High

3.2 : Pattern of regional inequality

3.2.1 Lorenz Curve and Gini Coefficient

Lorenz curve is a fairly widely used simple graphical method of comparing distributions on 2 - dimensional surfaces. Basically it uses a square graph with x - axis and y - axis having comparable scale, e.g., percent units (Fig.). Appropriate data are collected for subdivisions of the total area being considered. If the density of the objects distributed over the area is uniform, a straight-line curve results. It is called the line of equal distribution (LED) or 45° line with the following characteristics : (a) it is simply the SW - NE diagonal of the square graph, (b) it is obtained by joining the origin (0,0) of the graph and the point (100%, 100%), and (c) the slope of LED is + 45°. The degree of inequality of a distribution is directly proportional to the degree of concavity of the curve. Hence, the more the concavity, the more the inequality.



The Gini coefficient (G) is a measure of inequality of a distribution (Gini, 1912). It is defined as a ratio with boundary conditions, $0 = G = 1$. The numerator is the area between the Lorenz curve of the distribution and the uniform (perfect) distribution line (LED), while the denominator is the area under the uniform distribution line (LED). Here, $G = 0$ corresponds to perfect equality (i.e. each unit has the same quantity) and $G = 1$ corresponds to perfect inequality (i.e. only a single unit has all the quantity, while everyone else has zero). The Gini coefficient satisfies four important principles : *anonymity, scale independence, population independence, and transfer principle.* As a measure of inequality, G has several advantages – (a) it is a measure of inequality by means of a ratio analysis, (b) It can be used to compare distributions across different populations, (c) it is sufficiently simple and can be compared across countries and be easily interpreted, (d) it can be used to indicate how the distribution has changed within a country over a period of time and to see if inequality is increasing or decreasing. G can be calculated from the following formula :

$$G = 1 - \{Ó(x_i . y_{i+1}) - Ó(x_{i+1} . y_i)\}/10000$$

where, i = 1, 2, 3, ... n and (x, y) are co-ordinates of the points plotted

Example : Draw a Lorenz curve to show the nature of distribution of rural population of 10 Blocks and interpret.

Requirements :

1. A data base with total population and rural population of 10 Blocks
2. A Graph Paper (mm), Calculator, Pen, Pencil, Eraser etc.

Procedures :

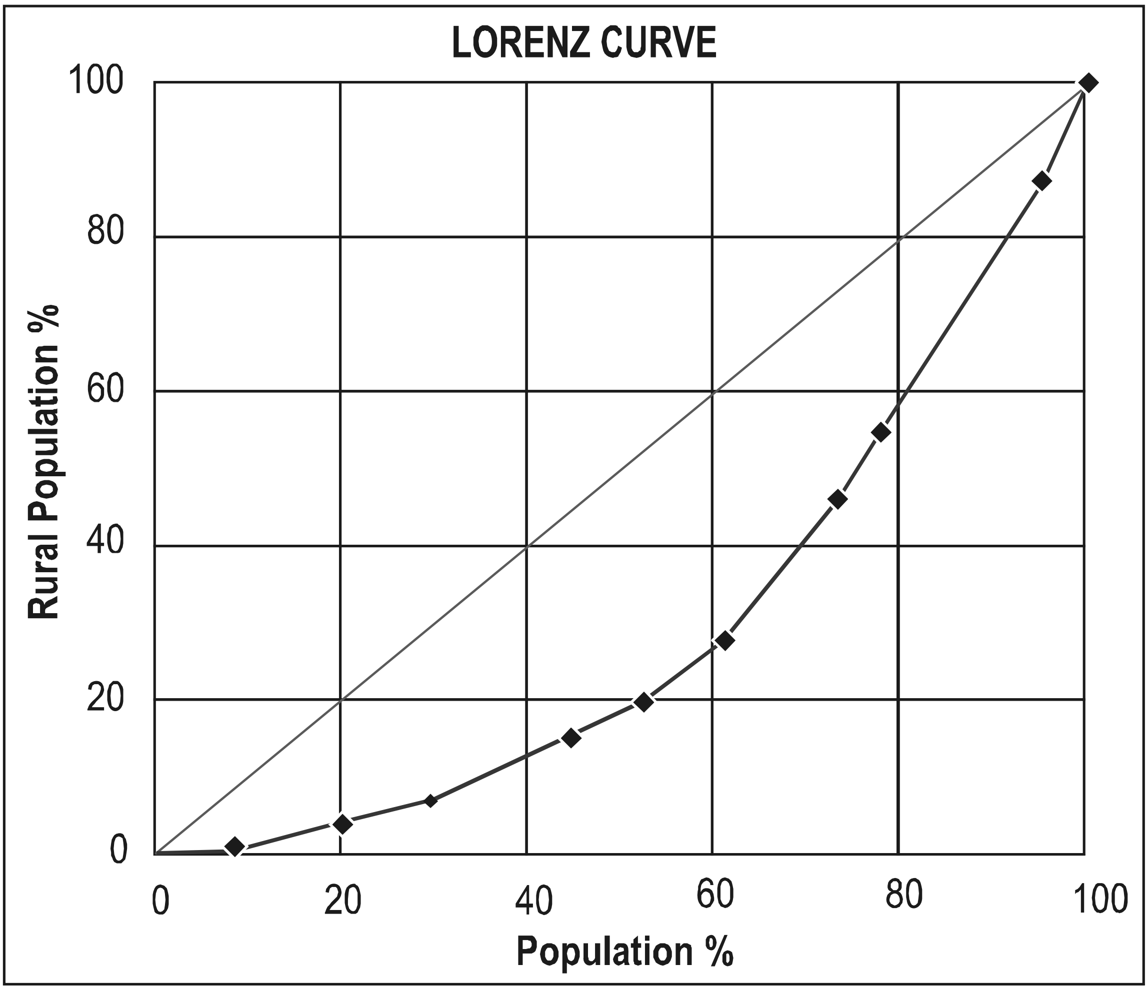
3. It is the 'rural population' of 10 blocks (Bi) of a district (D) for which Lorenz Curve is to be drawn
4. Compute the proportion of 'rural population' (Ri) of a block with respect to individual block total
5. Rank the Ri and redraw the Table according to the ascending rank order
6. Now compute the proportion of 'population' of a block with respect to the district total (BPi)
7. Also compute the proportion of 'rural population' of a block with respect to the district total (BRi)
8. Compute the cumulative frequencies of the BPi (x) and BRi (y) and enter in separate columns

Table – 7 : WORKSHEET FOR LORENZ CURVE

Blocks (Bi)	Population (Pi)	Rural Population (Ri)	Ri= (Ri/Pi) *100(%)	Rank based on % Ri	Redrawn Table			BPi= (Pi/DPi) *100(%)	Cumulative BPi (x:%)	BRi= (Ri/DRi) *100(%)	Cumulative BRi (y:%)
					Rank (ascending	Population (Pi)	Rural Population				
A	6050605	375033	6.19	7	1	3729644	10090	7.55	7.55	0.48	0.48
B	2555664	177501	6.94	8	2	5715030	70499	11.57	19.12	3.34	3.82
C	2805065	289906	10.33	10	3	4740149	61513	9.60	28.72	2.91	6.73
D	8331912	689636	8.28	9	4	7281881	169831	14.74	43.46	8.04	14.77
E	7281881	169831	2.33	4	5	3852097	90525	7.80	51.26	4.29	19.06
F	5715030	70499	1.23	2	6	4335230	176401	8.77	60.03	8.36	27.42
G	3729644	10090	0.27	1	7	6050605	375033	12.25	72.28	17.77	45.19
H	4335230	176401	4.07	6	8	2555664	177501	5.17	77.45	8.41	53.60
I	3852097	90525	2.35	5	9	8331912	689636	16.87	94.32	32.67	86.27
J	4740149	61513	1.29	3	10	2805065	289906	5.68	100.00	13.73	100.00
District Total	49397277	2110935				49397277	2110935	100.00		100.00	

Table – 7a : COMPUTATION OF GINI COEFFICIENT

Blocks	Xi	Yi	Xi . Yi+1	Xi+1 . Yi	Gini Coefficient
G	7.55	0.48	28.841	9.1776	G = 1 – {(25516.19 – 21477.09) / 10000} = 1 – {4039.094 / 10000} = 1 – 0.40 = 0.60
F	19.12	3.82	128.6776	109.7104	
J	28.72	6.73	424.1944	292.4858	
E	43.46	14.77	828.3476	757.1102	
I	51.26	19.06	1405.549	1144.172	
H	60.03	27.42	2712.756	1981.918	
A	72.28	45.19	3874.208	3499.966	
B	77.45	53.6	6681.612	5055.552	
D	94.32	86.27	9432	8627	
C	100	100	Sum = 25516.19	Sum = 21477.09	


LORENZ CURVE
100
80
60
40
20
0
Rural Population %
0
20
40
60
80
100
Population %

3.2. 2 : Z – score values

In statistics, a standard score (also called z-score or normal score) is a dimensionless quantity derived by subtracting the population mean from an individual (raw) score and then dividing the difference by the population standard deviation. It compares the various grading methods in a normal distribution. Two distributions with very different means and standard deviations are difficult to compare closely unless both distributions can be put into a standard form. The conversion process is called standardization. It considers both varying means and varying standard deviations.

The standard score is given by

The standard score is given by

$$z = \frac{X - \mu}{\sigma}$$

where

- X is a raw score to be standardized
- σ is the standard deviation of the population $(= \sqrt{\{\Sigma(x_i - x)^2 / N\}})$
- μ is the mean of the population $(= (\Sigma x_i)N)$
- N = number of districts

The quantity z represents the distance between the raw score and the population mean in units of the standard deviation. z is negative when the raw score is below the mean, positive when above. A key point is that calculating z requires the population mean and the population standard deviation, not the sample mean or sample deviation. It requires knowing the population statistics, not the statistics of a sample drawn from the population of interest.

Z – score values are very important in geographical analysis. It can be used in three ways : (a) to compare the nature of variations, particularly when attributes of more than one sample are compared, and as such to explore the degree of concentration or dispersion of attributes across samples, and (b) by way of this, it can well be used in classificatory problems. The more the value of Z – score, the more it is away from mean, or more it is distinct or specialized in a given pattern of distribution. On a Z – scale, samples may be easily placed for distributional class.

Example : Draw a Z – score map to show the nature of distribution of population of West Bengal, 2001

Requirements :

1. A data base with districtwise total population of West Bengal, 2001
2. Calculator, Pen, Pencil, Eraser etc.

Procedures :

1. Compute the arithmetic mean of total population of the districts
2. Compute the standard deviation of total population of the districts
3. Compute the Z – scores for each district
4. Prepare a distribution table with classes as : (0 ± 1), (1 ± 2), etc and draw choropleth maps and interpret

Note : Isopleth Map with isopleths of Z – scores, e.g., 0, 1, 2, 3, etc can also be drawn and interpreted

WORKSHEET FOR CALCULATION OF Z – SCORES

DISTRICTS	POPULATION 2001 (x)	PARAMETERS	$z = \frac{X - \mu}{\sigma}$
Purulia	2535233	$\mu = (\Sigma x)/N$	–0.80
Bankura	3191822	$= (80221171/18)$	–0.53
24-Paragana(N)	8930295	$= 4456732$ $\sigma = \sqrt{((x - \mu)^2/N)}$	1.87
Jalpaiguri	3403204	$= 2396161$	–0.44
Coochbehar	2478280		-0.83

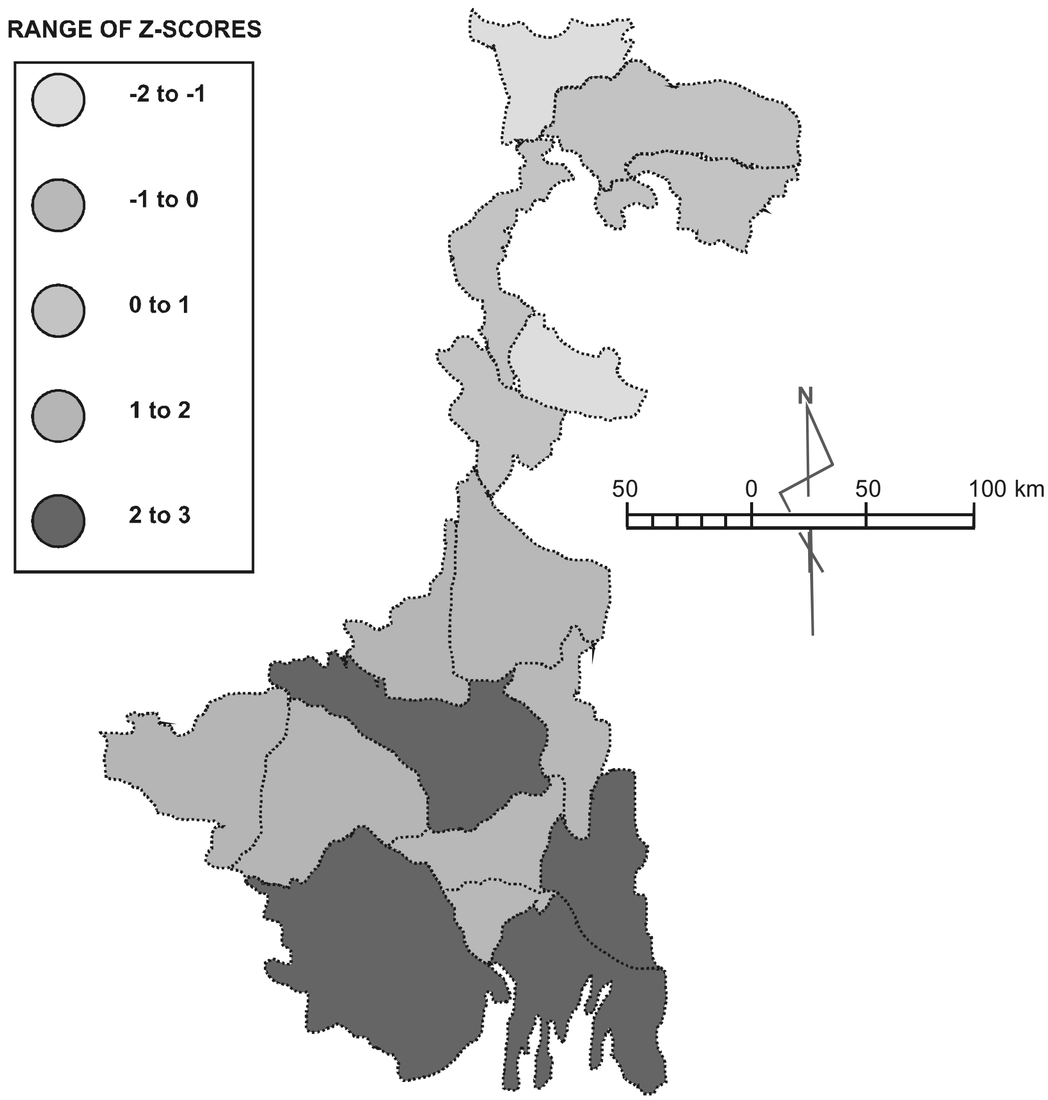
Table – 8 : WORKSHEET FOR Z – SCORES

DISTRICTS	POPULATION 2001 (x)	PARAMETERS	Z-SCORE $= (x - \overline{X})/\sigma$
Purulia	**2535233**		**-0.80**
Bankura	**3191822**	$\overline{X} = (\Sigma x)/N$	**-0.53**
Bankura	**3191822**	$= (80221171/18)$	**-0.53**
Midnapore	**9638473**	$= 4456732$	**2.16**
Birbhum	**3012546**	$\sigma = \sqrt{((x - \mu)^2/N)}$	**-0.60**
Burdwan	**6919698**	$= 2396161$	**1.03**
Nadia	**4603756**		**0.06**
U. Dinajpur	**2441824**		**-0.84**
Malda	**3290160**		**-0.49**
Hooghly	**5040047**		**0.24**
Howrah	**4274010**		**-0.08**
Murshidabad	**5863717**		**0.59**
Darjeeling	**1605900**		**-1.19**
Kolkata	**4580544**		**0.05**
D. Dinajpur	**1502647**		**-1.23**
24 Parganas (S)	**6909015**		**1.02**
24 Parganas (N)	**8930295**		**1.87**
Jalpaiguri	**3403204**		**-0.44**
Coochbehar	**2478280**		**-0.83**
Sum Total =	**80221171**		

Table – 8a : CHOROPLETH TABLE

Range of Z-Score	Districts	Remarks
-2 to 1	Darjeeling, D. Dinajpur	Least
-1 to 0	Purulia, Bankura, Birbhum, Howrah, Malda, U. Dinajpur, Jalpaiguri, Coochbehar	Less
0 to 1	Nadia, Hooghly, Kolkata, Murshidabad	Moderate
1 to 2	Burdwan, 24 Parganas (N), 24 Parganas (S)	High
2 to 3	Midnapore	Very High

DISTRIBUTION
OF
POPULATION OF WEST BENGAL, 2001
BY
Z-SCORE ANALYSIS



F - 22

Unit 4 □ Hierarchy Analysis

Structure

4.1. Rank – size distribution of towns

4.2. Functional Hierarchy of towns

4.1 Rank–size distribution of towns

When size is plotted against the rank for every town / city, the relationship (on a logarithmic scale) is shown by a downward sloping straight line (Fig.). This means that city size is inversely proportional to its rank. In other words, the product of the city size and its rank is constant, being equal to the size of the leading city in the system of cities (Auerbach, 1913; Zipf, 1949). If all the towns / cities of an area are ranked in descending order of population, the population of the nth ranked town will be one- nth that of the largest one.

$$\text{Thus, } P_r = P_1 \cdot (r)^{-q}$$

where r= rank of a city, P_r = population of a city of rank, r, P_1= population of the largest city and q=an exponent which generally has a value close to 1.

The tendency of the largest city to be "excessively" big with stunting effects on cities of nearby rank is the primate distribution case (Jefferson 1939, Linksy 1965, Harris 1971). According to Richardson (1973) the rank – size distribution may be interpreted as a very general model according to the value of the exponent : $q = 1$ implies the rank – size distribution, $q > 1$ represents metropolitan dominance, and $q < 1$ stands for an urban system in which intermediate cities are relatively large. The limiting cases are : $q =$ infinity (only one city) and $q = 0$ (all cities are of same size). The rank – size distribution matches well with the allometric growth model and Pareto distribution (Parr, 1970; Berry, 1961; Beckman, 1958; Nordbeck, 1971).

Example : Draw a rank – size graph for the 74 cities / towns of West Bengal, 2001

Requirements :

1. A data base with population of the 32 cities / towns of West Bengal, 2001
2. Log – log graph paper, Calculator, Pen, Pencil, Eraser etc.

Procedures :

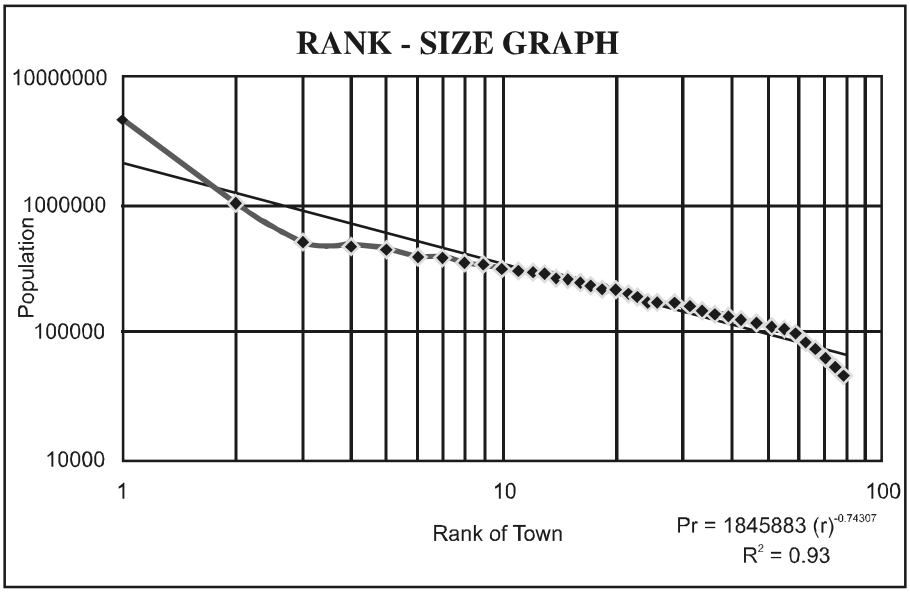
1. Arrange the cities / towns in descending order of their size of population
2. Assign ranks in ascending order so that the 1st ranked city has the largest population size

3. Take a 3 x 2 Cycle Log - log graph paper
4. Plot ranks along the x - axis and population along y - axis
5. Measure the map length of a cycle (d) from the graph in either cm or inch
6. Compute the perpendicular distance (away from the base of a cycle) for either population or rank using formula :

   y (cm or inch) = (log P - log Cb) x d

   x (cm or inch) = (log r - log Cb) x d

   where P = population of a city or town of rank r, Cb = Value of the base of a Cycle in which the required population or rank is to be plotted, and d = map length of a cycle
7. Plot the population of a city / town corresponding to their ranks and join them by lines.



(Fig-9)

Table – 9 : WORKSHEET FOR PLOTTING THE RANK – SIZE CURVE
(Log – Log Graph : Power Regression)

Rank (r)	Population (Pr)	log (r)	log (Pr)	sq. of log (r)	log (r) x log (Pr)
1	4580544	0	6.660917059	0	0
2	1008704	0.301	6.003763743	0.090619058	1.807312974
3	492996	0.4771	5.692843396	0.227644692	2.716176584
4	486304	0.6021	5.686907842	0.362476233	3.423859686
5	441956	0.699	5.645379034	0.488559067	3.945950608
6	392150	0.7782	5.593452219	0.605519368	4.352551838
7	389214	0.8451	5.590188453	0.714190697	4.724257305
8	348379	0.9031	5.542051968	0.815571525	5.00497164
9	336390	0.9542	5.526843077	0.910578767	5.273948607
10	314334	1	5.497391359	1	5.497391359
11	290067	1.0414	5.462498323	1.084498725	5.688605797
12	285871	1.0792	5.456170101	1.164632162	5.888196448
13	284615	1.1139	5.454257785	1.240869792	6.075734201
14	271781	1.1461	5.434219092	1.313609474	6.228310854
15	261575	1.1761	5.417596234	1.38319065	6.371587576
16	250615	1.2041	5.399007061	1.449904933	6.501052289
17	231515	1.2304	5.364579134	1.514004548	6.60084061
18	220032	1.2553	5.342485846	1.575709062	6.706275592
19	215432	1.2788	5.333310213	1.635210772	6.81998964
20	207984	1.301	5.318029926	1.69267905	6.918916452
21	202095	1.3222	5.305555569	1.748263863	7.015107942

Solution of the Rank – Size Equation : $P_r = P_1 (r)^b$

The normal equations are :
?log P = N. a + b ?log r
?(logP. Log r) = a . ?log r + b ?(log r)^2
From Table :
?log P = 174.1988872
?log r = 35.42
?(logP. Log r) = 189.6336701
?(log r)^2 = 43.48870256
Thus,
174.1988872 = 32 a + 35.42 b
189.6336701 = 35.42 b + 43.48870256 b

Eliminating coefficients of ' a ',
5.443715225 = a + 1.106875 b
5.353858557 = a + 1.122780075 b

Subtracting and transposing,
b = – 0. 743073
a = 6.26620417

P_1 = antilog (6.26620417)
= 1845883

Therefore, $P_r = 1845883 (r)^{-0.743073}$

Table – 9 : WORKSHEET FOR PLOTTING THE RANK – SIZE CURVE *(Contd.)*

Rank (r)	Population (Pr)	log (r)	log (Pr)	sq. of log (r)	log (r) x log (Pr)	Solution of the Rank – Size Equation : $P_r = P_1 (r)^b$
22	197955	1.3424	5.296566476	1.802098654	7.110230968	
23	185660	1.3617	5.268718346	1.854302699	7.174560432	
24	170695	1.3802	5.2322208	1.904983072	7.221569967	
25	170201	1.3979	5.230962107	1.954236268	7.312571214	
26	167848	1.415	5.224916171	2.002149575	7.393117127	
27	165222	1.4314	5.218067875	2.048802225	7.468953275	
28	162166	1.4472	5.209959804	2.094266368	7.539635174	
29	161448	1.4624	5.208032669	2.138607904	7.616216549	
30	160168	1.4771	5.204575753	2.181887201	7.687789466	
31	155503	1.4914	5.191738772	2.224159702	7.742760329	
32	153349	1.5051	5.185680948	2.265476457	7.805227567	
	Total =	35.42	174.1988872	43.48870256	189.6336701	

Table – 9a : WORKSHEET FOR PLOTTING THE RANK – SIZE CURVE (Arithmetic Graph)

Rank (r)	Population (Pr)	1 / (r)	Computation of Estimated Population of the 1st ranked Town	Computation of Estimated Population of Towns $Pr = P_1 / r$
1	4580544	1	Estimated Population of the 1st Ranked town / city is given by –	329254
2	1008704	0.5		1646271
3	492996	0.333333333		1097514
4	486304	0.25	P_1 = ?(Pr) / ?(1 / r)	823135
5	441956	0.2	= 13362768 / 4.0585	658508
6	392150	0.166666667	= 3292543	548757
7	389214	0.142857143		470363

Table – 9a : WORKSHEET FOR PLOTTING THE RANK – SIZE CURVE *(Contd.)*

Rank (r)	Population (Pr)	1 / (r)	Computation of Estimated Population of the 1st ranked Town	Computation of Estimated Population of Towns $Pr = P_1 / r$
8	348379	0.125		411567
9	336390	0.111111111		365838
10	314334	0.1		329254
11	290067	0.090909091		299322
12	285871	0.083333333		274378
13	284615	0.076923077		253272
14	271781	0.071428571		235182
15	261575	0.066666667		219503
16	250615	0.0625		205784
17	231515	0.058823529		193679
18	220032	0.055555556		182919
19	215432	0.052631579		173292
20	207984	0.05		164627
21	202095	0.047619048		156787
22	197955	0.045454545		149661
23	185660	0.043478261		143154
24	170695	0.041666667		137189
25	170201	0.04		131702
26	167848	0.038461538		126636
27	165222	0.037037037		121946
28	162166	0.035714286		117591
29	161448	0.034482759		113535
30	160168	0.033333333		109751
31	155503	0.032258065		106211
32	153349	0.03125		102892
	13362768	4.058495195		

4.2 : Functional Hierarchy of Towns

As aggregates of human population, towns are devoted to a number of functions, performed by the working section of their inhabitants (Harris, 1943). The number, relative proportion and character of these functions indicate the nature of urbanization in a particular environmental setting. The pattern of functions helps to delineate the regional system of towns/cities along with the hierarchy. Harris (1943) used the proportion of labour force in a particular occupation as the basic criteria for determining the degree of its specialization. Pownall (1953) later modified the scheme by using 'mean' and 'deviation from mean' in finding the 'distinctiveness' of a function. Later Nelson (1955) used the 'mean' and 'standard deviation' to determine the 'degree of distinctiveness' of any function. A town which shows a percentage employment of more than 'mean' plus one 'standard deviation' is said to be significantly characterized by the function diagnosed by the occupation group. This is further developed by recording how many times the employment ratio in one town is above the 'mean' for all towns in terms of the 'standard deviation' (Dick, 1961; Dacey, 1962; Hadden and Borgatta, 1965; Berry, 1972; Ram and Sinha, 1972; Mitra, 1981).

Example : Classify the following towns according to their functions

Requirements :

1. A town data base with occupational pattern of population
2. Graph paper, Calculator, Pen, Pencil, Eraser etc.

Procedures :

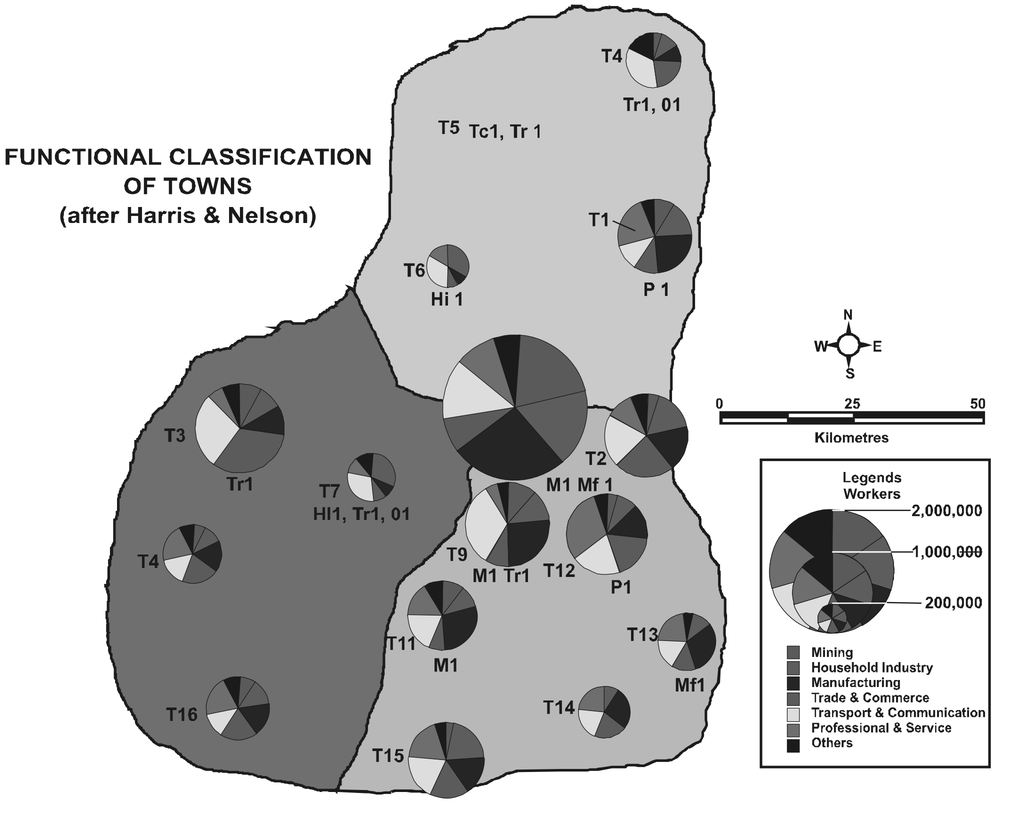
1. Prepare a data matrix with the following columns as shown below –

Towns	Workers (%)			
	Occupation – 1	Occupation – 2	Occupation – 3	Occupation – n
AAA				
BBB				
Mean				
Standard Deviation				

2. Compute the 'mean' of the workers (%) separately for all the occupations
3. Compute the 'standard deviation' of the workers (%) separately for all the occupations
4. Draw frequency curves to show the distribution of workers in different occupations
5. For each occupation prepare separate classificatory tables as shown below –

Occupation – 1	Degree of Specialization	Occupation – 2	Degree of Specialization
Mean to (Mean + 1 SD)		Mean to (Mean + 1 SD)	
(Mean + 1SD) to (Mean + 2SD)	$O1_1$	(Mean + 1SD) to (Mean + 2SD)	$O2_1$
(Mean + 2SD) to (Mean + 3SD)	$O1_2$	(Mean + 2SD) to (Mean + 3SD)	$O2_2$
(Mean + 3SD) to (Mean + 4SD)	$O1_3$	(Mean + 3SD) to (Mean + 4SD)	$O2_3$

6. Assign the 'degree of functional specilisation' to the Towns to complete the process of classification and identification of hierarchy.



(Fig-10)

Table – 10 : WORKSHEET FOR FUNCTIONAL CLASSIFICATION OF TOWNS / CITIES

Towns	Proportion of Workers employed (%)							Nomenclature of Towns
	Mining (M)	Household Industry (HI)	Manufacturing Industry (Mf)	Trade & Commerce (Tc)	Transport & Communication (Tr)	Professional Services (P)	Others (O)	
T1	8.8	16.3	22.2	10.8	12.4	**25.2 (1)**	4.3	P_1 : Service Town
T2	4.5	18.3	16.7	22.3	20.8	12.3	5.1	*Diversified*
T3	7.8	9.8	10.9	**31.6 (1)**	27.4	7.7	4.8	Tc_1 : Trade & Commerce Town
T4	3.9	12.4	8.6	22.9	**34.2 (1)**	9.8	**8.2 (1)**	Tr_1 O_1: Transport Town and Others
T5	2.1	9.6	15.8	**31.8 (1)**	18.7	11.5	**10.5 (1)**	Tc_1 Tr_1: Trade & Commerce and Transport Town
T6	0	**34.2 (2)**	8.2	7.2	**31.2 (1)**	18.2	1.0	HI_1 : Household Industry
T7	0	**32.6 (2)**	6.4	8.5	**30.7 (1)**	11.9	**9.9 (1)**	HI_1 Tr_1 O_1: Household Industry, Transport Town and Others
T8	5.6	9.9	18.4	22.1	18.2	19.6	6.2	*Diversified*
T9	**12.3 (1)**	11.3	24.5	9.5	**34.2 (1)**	6.1	2.1	M_1 Tr_1: Mining and Transport Town
T10	**21.1 (2)**	17.2	**26.9 (1)**	7.3	12.7	11.1	3.7	M_1 Mf_1: Mining and Manufacturing Town
T11	9.7	10.6	**28.2 (1)**	6.9	18.9	18.4	7.3	Mf_1: Manufacturing Town
T12	3.4	8.6	14.4	18.4	20.5	**30.5 (1)**	4.2	P_1: Service Town
T13	2.1	12.7	**30.8 (1)**	12.8	17.2	23.1	1.3	Mf_1: Manufacturing Town
T14	2.5	5.8	23.8	24.4	20.2	22.4	0.9	*Diversified*
T15	3.9	21.5	14.6	16.5	21.1	18.6	3.8	*Diversified*
T16	8.3	15.2	15.4	21.2	12.3	20.6	7	*Diversified*
Mean	6	15.4	17.9	17.1	21.9	16.7	5.0	
Standard Deviation	5.4	8.1	7.5	8.5	7.4	6.9	3.0	