

TABLE OF CONTENTS

1.	THE CLUSTER PROBLEM AND PRELIMINARY IDEAS	1
1.1	Basic Notions and Definitions	1
1.2	The Cluster Problem	1
1.3	Distance Functions	3
1.4	Measures of Similarity	5
1.5	Distance and Similarity Between Clusters	11
1.6	Cluster Methods Based on Euclidean Distance	19
1.7	An Algorithm for Hierarchical Clustering	24
1.8	Other Aspects of the Cluster Problem	30
2.	CLUSTERING BY COMPLETE ENUMERATION	32
2.1	Introduction	32
2.2	The Number of Partitions of n Objects into m Non-empty Subsets	32
2.3	Recursive Relation for Stirling's Numbers of the Second Kind	39
2.4	Computational Aspects of Complete Enumeration	41
3.	MATHEMATICAL PROGRAMMING AND CLUSTER ANALYSIS	43
3.1	Application of Dynamic Programming to the Cluster Problem	43
3.2	Jensen's Dynamic Programming Model	52
3.3	Integer Programming Applications to Cluster Analysis	63
4.	SIMILARITY MATRIX REPRESENTATIONS	73
4.1	Dendograms	73
4.2	Comparison of Dendograms or Their Similarity Matrices	79
4.3	Basic Definitions	80

4.4	Trees	82
4.5	Local Operations on Trees	88
5.	CLUSTERING BASED ON DENSITY ESTIMATION	91
5.1	Mode Analysis	91
5.2	Probability Density Function Estimation	93
5.3	Clustering Based on Density Estimation	96
5.4	Remarks	99
6.	APPLICATIONS	100
6.1	Application to Remote Sensing Data	100
6.2	Application of Density Estimation Technique to Fisher's Iris Data	103
7.	HISTORICAL COMMENTS	104
	REFERENCES	110