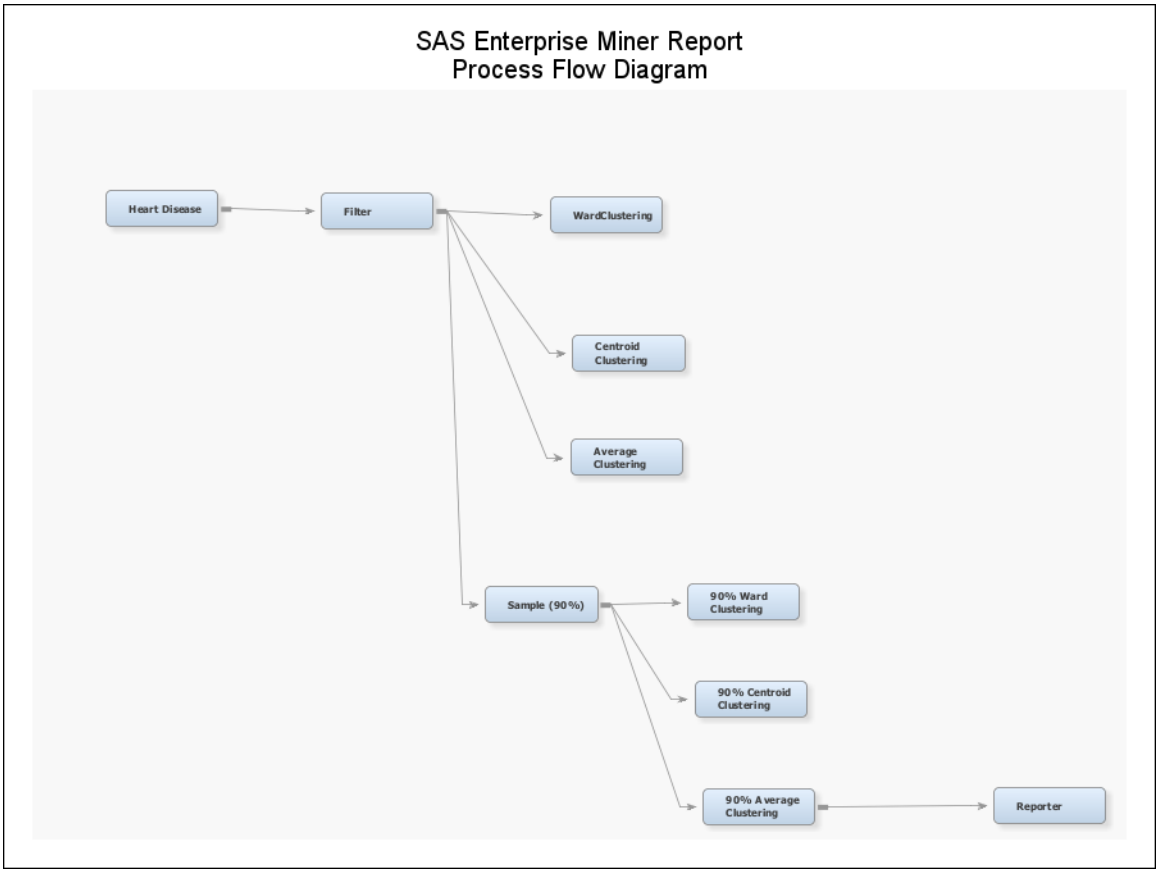
SAS Enterprise Miner Report

User = gmark1
Date = 22:59:23 May 04
Project = Heart disease
Diagram = GK.Cluster Analysis

Start Node = Report
Node label = Reporter
Nodes = PATH
Showall = N

Format = PDF
Style = LISTING



SAS Enterprise Miner Report

Node=Heart Disease
Summary

Node id = Ids2
Node label = Heart Disease
Meta path = Ids2
Notes =

Node=Heart Disease
Properties

Property	Value	Default	Property	Value	Default	Property	Value	Default
Component	DataSource		DsCreatedBy	gmark1		NBytes	4391936	.
ApplyIntervalLevelLowerLimit	Y		DsId	emsavetrain		NCols	18	.
ApplyMaxClassLevels	Y		DsModifiedBy	gmark1		NObs	29944	.
ApplyMaxPercentMissing	Y		DsModifyDate	2030480026.2		NewTable		
CMeta	WORK.M2W9XXHV		DsSampleName			NewVariableRole	REJECT	
ComputeStatistics	N		DsSampleSize			OutputType	VIEW	
DBPassThrough	Y		DsSampleSizeType			Role	RAW	TRAIN
Data	GKDATA.EM_SAVE_TRAIN		DsScope	LOCAL		Sample	D	
DataSelection	DATASOURCE		IdentifyEmptyColumns	Y		SampleSizeObs	10000	
DataSource	emsavetrain		IntervalLowerLimit	20		SampleSizePercent	20	
DataSourceRole	RAW		Library	GKDATA		SampleSizeType	PERCENT	
Description			MaxClassLevels	20		Scope	LOCAL	
DropMapVariables	Y		MaxPercentMissing	50		Segment		
DsCreateDate	2030480025.9		MetaAdvisor	BASIC		Table	EM_SAVE_TRAIN	

Node=Heart Disease
Data Attributes

Attribute	Value	Attribute	Value	Attribute	Value
Data Name	EM_SAVE_TRAIN	Date Created	04May2024:22:12:17	Data Size	4391936
Data Type	DATA	Date Modified	04May2024:22:12:17	Role	RAW
Data Label		Number Rows	29944	Segment	
Engine	BASE	Number Columns	18	Data Library	GKDATA

Node=Heart Disease
Variables List

Name	Label	Role	Level	Type	Length	Format	Creator
AgeCategory	AgeCategory	INPUT	NOMINAL	C	11	$11.	
AlcoholDrinking	AlcoholDrinking	INPUT	NOMINAL	C	3	$3.	
Asthma	Asthma	INPUT	NOMINAL	C	3	$3.	
BMI	BMI	INPUT	INTERVAL	N	8	BEST.	
Diabetic	Diabetic	INPUT	NOMINAL	C	23	$23.	
Difficult_Walking		INPUT	NOMINAL	C	3		
General_Health		INPUT	NOMINAL	C	9		
HeartDisease	HeartDisease	INPUT	NOMINAL	C	3	$3.	
KidneyDisease	KidneyDisease	INPUT	NOMINAL	C	3	$3.	
MentalHealth	MentalHealth	INPUT	INTERVAL	N	8	BEST.	
PhysicalActivity	PhysicalActivity	INPUT	NOMINAL	C	3	$3.	

Name	Label	Role	Level	Type	Length	Format	Creator
PhysicalHealth	PhysicalHealth	INPUT	INTERVAL	N	8	BEST.	
Race	Race	INPUT	NOMINAL	C	30	$30.	
Sex	Sex	INPUT	NOMINAL	C	6	$6.	
SkinCancer	SkinCancer	INPUT	NOMINAL	C	3	$3.	
SleepTime	SleepTime	INPUT	INTERVAL	N	8	BEST.	
Smoking	Smoking	INPUT	NOMINAL	C	3	$3.	
Stroke	Stroke	INPUT	NOMINAL	C	3	$3.	

SAS Enterprise Miner Report

Node=Filter
Summary

Node id = Filter
Node label = Filter
Meta path = lds2 => Filter
Notes =

Node=Filter
Properties

Property	Value	Default	Property	Value	Default	Property	Value	Default
Component	Filter		KeepMissingInterval	Y		PercentsCutoff	0.5	
ClassFilterMethod	MINPCT		MADSCutoff	9		PublishScoreCode	Y	
CreateDistributionData	Y	N	MaxValues	25		SpacingsCutoff	9	
ExportTable	FILTERED		MinFreq	1		StddevCutoff	3	
IntervalFilterMethod	STDDEV		MinPercent	0.01		TabletoFilter	TRAIN	
KeepMissingClass	Y		NormalizeClassValue	Y		UpdateClassLevel	N	

Node=Filter
Variable Summary

Role	Level	Frequency Count	Name
INPUT	INTERVAL	4	BMI MentalHealth PhysicalHealth SleepTime
INPUT	NOMINAL	14	AgeCategory AlcoholDrinking Asthma Diabetic Difficult_Walking General_Health HeartDisease KidneyDisease PhysicalActivity Race Sex SkinCancer Smoking Stroke

Node=Filter
Excluded Class Values

Variable	Role	Level	Train Count	Train Percent	Label	Filter Method
Diabetic	INPUT	YES (DURING PREGNANCY)	177	0.59110	Diabetic	MINPCT

Node=Filter
Limits for Interval Variables

Variable	Role	Minimum	Maximum	Filter Method	Keep Missing Values	Label
BMI	INPUT	9.2635	48.6838	STDDEV	Y	BMI
MentalHealth	INPUT	-22.3426	31.5833	STDDEV	Y	MentalHealth
PhysicalHealth	INPUT	-25.7158	39.0751	STDDEV	Y	PhysicalHealth
SleepTime	INPUT	1.8932	12.4110	STDDEV	Y	SleepTime

SAS Enterprise Miner Report

Node=Sample (90%)
Summary

Node id = Smpl
Node label = Sample (90%)
Meta path = Ids2 => Filter => Smpl
Notes =

Node=Sample (90%)
Properties

Property	Value	Default	Property	Value	Default	Property	Value	Default
Component	Sample		IgnoreSmallStrata	N		OutputType	DATA	
AdjustFreq	N		IntervalDistribution	Y		Pvalue	0.01	
Alpha	0.01		LevelProportion	100		RandomSeed	12345	
ClassDistribution	Y		LevelSampleProportion	50		SizeObs	.	
ClusterMethod	RANDOM		LevelSelection	EVENT		SizePercent	10	
FreqCount	N		Method	DEFAULT		SizeType	PERCENT	
FreqMiss	N		MinStrataSize	5		StratifyCriterion	PROPORTIONAL	

Node=Sample (90%)
Variable Summary

Role	Level	Frequency Count	Name
INPUT	INTERVAL	4	BMI MentalHealth PhysicalHealth SleepTime
INPUT	NOMINAL	14	AgeCategory AlcoholDrinking Asthma Diabetic Difficult_Walking General_Health HeartDisease KidneyDisease PhysicalActivity Race Sex SkinCancer Smoking Stroke

SAS Enterprise Miner Report

Node=90% Average Clustering Summary

Node id = Clus6
Node label = 90% Average Clustering
Meta path = Ids2 => Filter => Smpl => Clus6
Notes =

Node=90% Average Clustering Properties

Property	Value	Default	Property	Value	Default	Property	Value	Default
Component	Cluster		FinalMaxNum	20		MissingNominal	DEFAULT	
AutomaticMaxNum	50		HideVariable	Y		MissingOrdinal	DEFAULT	
AutomaticMethod	CENTROID	WARD	ImputationMethod	NONE		NominalEncoding	GLM	
AutomaticMinNum	2		Initial	DEFAULT		NumberClusterMethod	AUTOMATIC	
Bins	100		Learn	.		OrdinalEncoding	RANK	
CCCCutOff	3		LearnFinal	0.02		Radius	0	
ClusterGraphs	Y		LearnInitial	0.5		Stdize	STD	
Clusvar	_SEGMENT_		LearnSteps	1000		Summary	N	
ClusvarLabel	Segment Variable		MaxC	10		TrainDefaults	Y	
ClusvarRole	SEGMENT		MaxIter	10		TreeProfile	Y	
DistancePlot	Y		MaxSteps	1200		XConv	0.0001	
Drift	N		MissingInterval	DEFAULT				

Node=90% Average Clustering Variable Summary

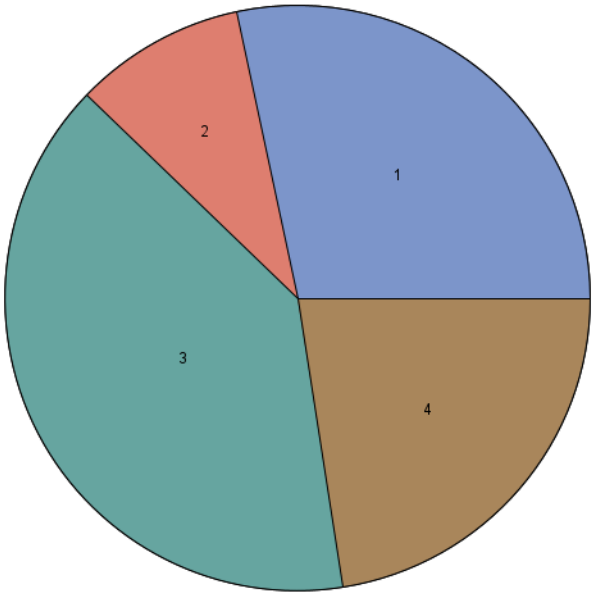
Role	Level	Frequency Count	Name
INPUT	INTERVAL	4	BMI MentalHealth PhysicalHealth SleepTime
INPUT	NOMINAL	14	AgeCategory AlcoholDrinking Asthma Diabetic Difficult_Walking General_Health HeartDisease KidneyDisease PhysicalActivity Race Sex SkinCancer Smoking Stroke
ID	INTERVAL	1	_dataobs_

Node=90% Average Clustering Mean Statistics

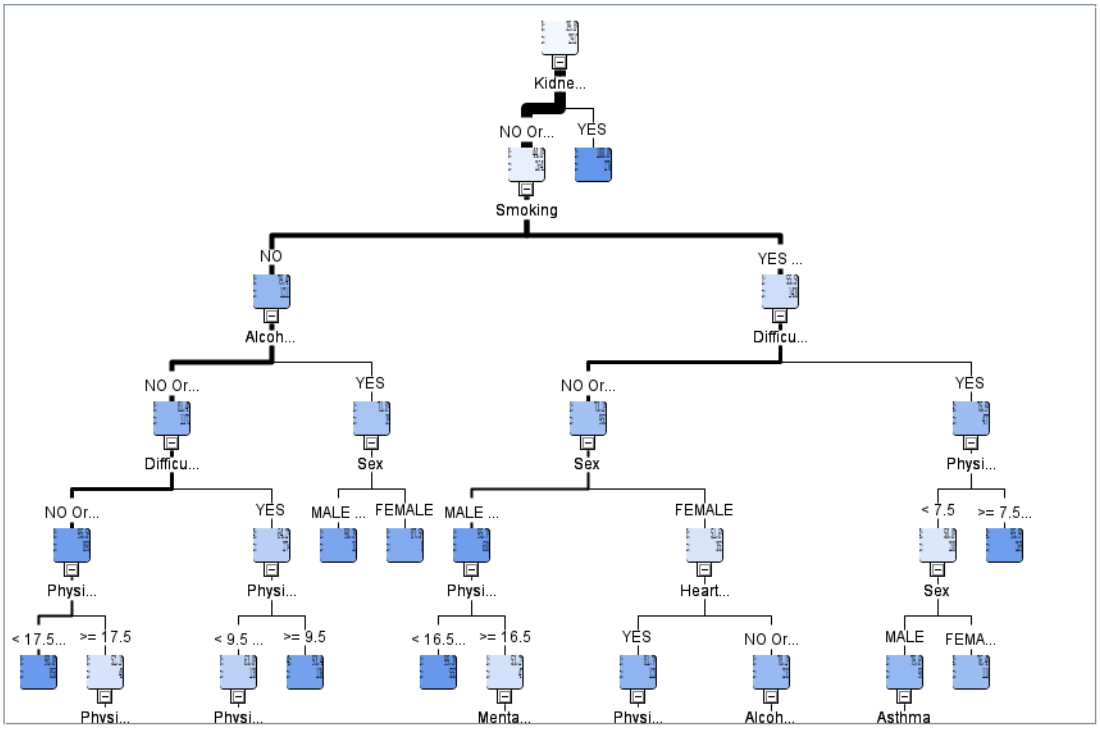
Variable	Segment 1	Segment 2	Segment 3	Segment 4
Clustering Criterion	0.536	0.536	0.54	0.536
Maximum Relative Change in Cluster Seeds	0.003	0.003	0.00	0.003
Improvement in Clustering Criterion	.	.	.	.
Frequency of Cluster	826.000	276.000	1162.00	653.000
Root-Mean-Square Standard Deviation	0.520	0.600	0.49	0.603
Maximum Distance from Cluster Seed	7.045	7.474	6.62	7.530
Nearest Cluster	3.000	4.000	1.00	1.000
Distance to Nearest Cluster	2.023	3.682	2.02	3.098
BMI	28.381	29.157	27.75	30.346
MentalHealth	2.196	5.395	2.39	11.009
PhysicalHealth	2.530	11.764	1.77	18.515
SleepTime	7.237	6.899	7.31	6.504
AgeCategory=18-24	0.019	0.000	0.05	-0.000

Variable	Segment 1	Segment 2	Segment 3	Segment 4
AgeCategory=25-29	0.019	0.014	0.02	0.002
AgeCategory=30-34	0.040	0.011	0.04	0.005
AgeCategory=35-39	0.038	0.007	0.03	0.025
AgeCategory=40-44	0.029	0.014	0.04	0.021
AgeCategory=45-49	0.034	0.029	0.06	0.040
AgeCategory=50-54	0.042	0.029	0.05	0.086
AgeCategory=55-59	0.061	0.062	0.07	0.116
AgeCategory=60-64	0.115	0.109	0.10	0.135
AgeCategory=65-69	0.150	0.116	0.13	0.136
AgeCategory=70-74	0.180	0.181	0.13	0.141
AgeCategory=75-79	0.125	0.159	0.13	0.141
AgeCategory=80 or older	0.148	0.268	0.14	0.153
AlcoholDrinking=No	0.878	0.978	0.99	0.971
AlcoholDrinking=Yes	0.122	0.022	0.01	0.029
Asthma=No	0.926	0.779	0.89	0.671
Asthma=Yes	0.074	0.221	0.11	0.329
Diabetic=No	0.752	0.478	0.85	0.560
Diabetic=No, borderline diabetes	0.025	0.022	0.02	0.029
Diabetic=Yes	0.223	0.500	0.13	0.410
Difficult_Walking=No	0.877	0.464	0.90	0.173
Difficult_Walking=Yes	0.123	0.536	0.10	0.827
General_Health=Excellent	0.090	0.022	0.16	0.012
General_Health=Fair	0.185	0.293	0.11	0.424
General_Health=Good	0.414	0.322	0.35	0.211
General_Health=Poor	0.021	0.246	0.01	0.312
General_Health=Very good	0.291	0.116	0.36	0.040
HeartDisease=No	0.424	0.330	0.72	0.391
HeartDisease=Yes	0.576	0.670	0.28	0.609
KidneyDisease=No	1.000	0.000	1.00	1.000
KidneyDisease=Yes	-0.000	1.000	-0.00	-0.000
PhysicalActivity=No	0.234	0.471	0.17	0.663
PhysicalActivity=Yes	0.766	0.529	0.83	0.337
Race=American Indian/Alaskan Native	0.030	0.011	0.03	0.032
Race=Asian	0.005	0.007	0.02	0.011
Race=Black	0.065	0.094	0.10	0.121
Race=Hispanic	0.051	0.040	0.05	0.032
Race=Other	0.035	0.029	0.03	0.035
Race=White	0.814	0.819	0.76	0.769
Sex=Female	0.161	0.612	0.60	0.658
Sex=Male	0.839	0.388	0.40	0.342
SkinCancer=No	0.841	0.714	0.86	0.864
SkinCancer=Yes	0.159	0.286	0.14	0.136
Smoking=No	0.040	0.435	0.83	0.332
Smoking=Yes	0.960	0.565	0.17	0.668
Stroke=No	0.768	0.514	0.69	0.461
Stroke=Yes	0.232	0.486	0.31	0.539

SAS Enterprise Miner Report
Node=90% Average Clustering
Segment Size



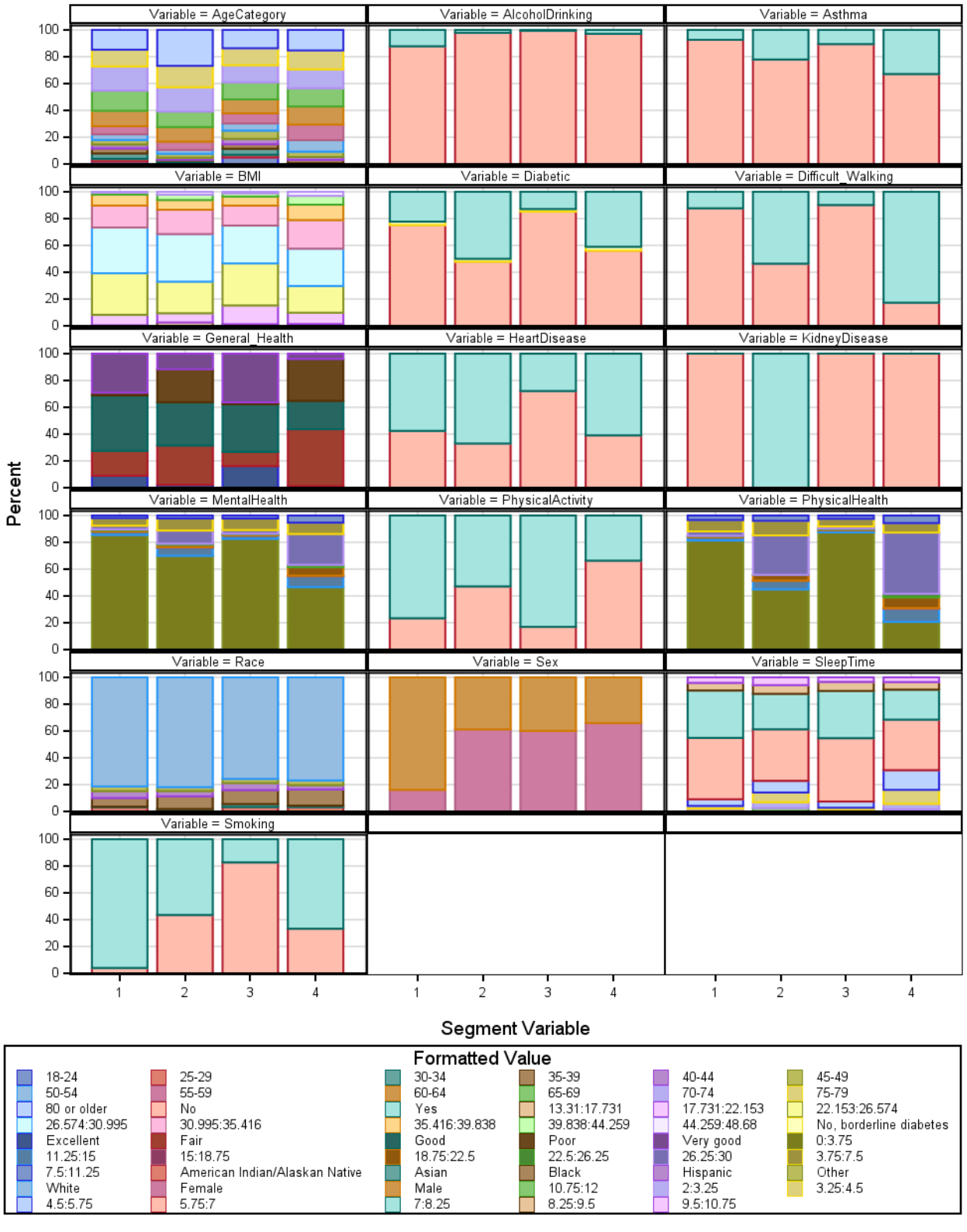
SAS Enterprise Miner Report
Node=90% Average Clustering
Tree Diagram

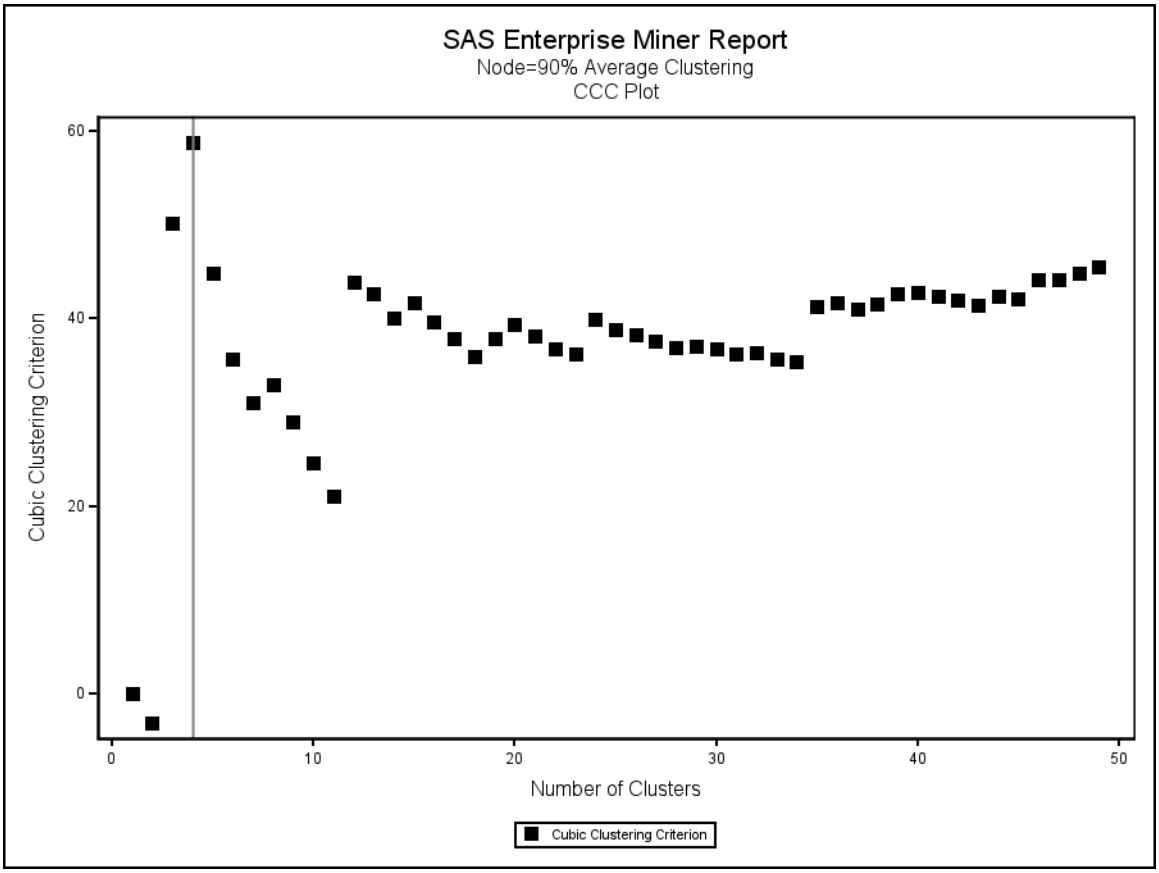


SAS Enterprise Miner Report

Node=90% Average Clustering

Segment Plot





End of Report