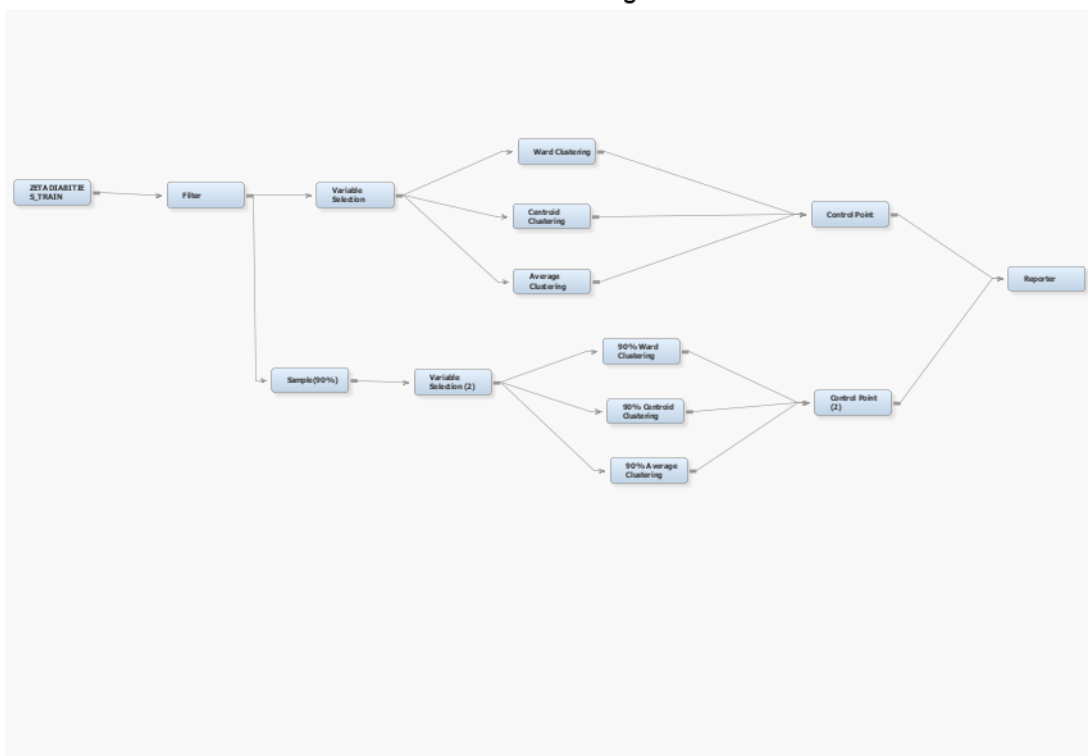# SAS Enterprise Miner Report

User   = bkris2
Date   = 15:49:10 March 09
Project   = ZetaExploratoryAnalysis
Diagram = ZetaExploratoryAnalysis

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

Format   = PDF
Style   = JOURNAL

SAS Enterprise Miner Report
Process Flow Diagram

SAS Enterprise Miner Report

Node=ZETADIABITIES_TRAIN
Summary

Node id = lds
Node label = ZETADIABITIES_TRAIN
Meta path = lds
Notes =

Node=ZETADIABITIES_TRAIN
Properties

Property	Value	Default	Property	Value	Default	Property	Value	Default
Component	DataSource		DsCreatedBy	bkris2		NBytes	44303360	.
ApplyIntervalLevelLowerLimit	Y		DsId	zetadiabitiestrain		NCols	22	.
ApplyMaxClassLevels	Y		DsModifiedBy	bkris2		NObs	250716	.
ApplyMaxPercentMissing	Y		DsModifyDate	2025037601.1		NewTable		
CMeta	WORK.M3LPL574		DsSampleName			NewVariableRole	REJECT	
ComputeStatistics	N		DsSampleSize			OutputType	VIEW	
DBPassThrough	Y		DsSampleSizeType			Role	RAW	TRAIN
Data	ZETADATA.ZETADIABITIES_TRAIN		DsScope	LOCAL		Sample	D	
DataSelection	DATASOURCE		IdentifyEmptyColumns	Y		SampleSizeObs	10000	
DataSource	zetadiabitiestrain		IntervalLowerLimit	20		SampleSizePercent	20	
DataSourceRole	RAW		Library	ZETADATA		SampleSizeType	PERCENT	
Description			MaxClassLevels	20		Scope	LOCAL	
DropMapVariables	Y		MaxPercentMissing	50		Segment		
DsCreateDate	2025037600.8		MetaAdvisor	BASIC		Table	ZETADIABITIES_TRAIN	

Node=ZETADIABITIES_TRAIN
Data Attributes

Attribute	Value	Attribute	Value	Attribute	Value
Data Name	ZETADIABITIES_TRAIN	Date Created	02Mar2024:22:24:48	Data Size	44303360
Data Type	DATA	Date Modified	02Mar2024:22:24:48	Role	RAW
Data Label		Number Rows	250716	Segment	
Engine	BASE	Number Columns	22	Data Library	ZETADATA

Node=ZETADIABITIES_TRAIN
Variables List

Name	Label	Role	Level	Type	Length	Format	Creator
Age		INPUT	INTERVAL	N	8	BEST12.0	
AnyHealthcare		INPUT	BINARY	N	8	BEST12.0	
BMI		INPUT	INTERVAL	N	8	BEST12.0	
CholCheck		INPUT	BINARY	N	8	BEST12.0	
Diabetes_012		REJECTED	NOMINAL	N	8	BEST12.0	
DiffWalk		INPUT	BINARY	N	8	BEST12.0	
Education		INPUT	ORDINAL	N	8	BEST12.0	
Fruits		INPUT	BINARY	N	8	BEST12.0	
GenHlth		INPUT	ORDINAL	N	8	BEST12.0	
HeartDiseaseorAttack		INPUT	BINARY	N	8	BEST12.0	
HighBP		INPUT	BINARY	N	8	BEST12.0	

Name	Label	Role	Level	Type	Length	Format	Creator
HighChol		INPUT	BINARY	N	8	BEST12.0	
HvyAlcoholConsump		INPUT	BINARY	N	8	BEST12.0	
Income		INPUT	ORDINAL	N	8	BEST12.0	
MentHlth		INPUT	ORDINAL	N	8	BEST12.0	
NoDocbcCost		INPUT	BINARY	N	8	BEST12.0	
PhysActivity		INPUT	BINARY	N	8	BEST12.0	
PhysHlth		INPUT	ORDINAL	N	8	BEST12.0	
Sex		INPUT	BINARY	N	8	BEST12.0	
Smoker		INPUT	BINARY	N	8	BEST12.0	
Stroke		INPUT	BINARY	N	8	BEST12.0	
Veggies		INPUT	BINARY	N	8	BEST12.0	

SAS Enterprise Miner Report

Node=Filter
Summary

Node id = Filter
Node label = Filter
Meta path = Ids => 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
REJECTED	NOMINAL	1	Diabetes_012
INPUT	BINARY	14	AnyHealthcare CholCheck DiffWalk Fruits HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost PhysActivity Sex Smoker Stroke Veggies
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth

Node=Filter
Excluded Class Values

Variable	Role	Level	Train Count	Train Percent	Label	Filter Method
Education	INPUT	1	167	0.066609		MINPCT

Node=Filter
Limits for Interval Variables

Variable	Role	Minimum	Maximum	Filter Method	Keep Missing Values	Label
Age	INPUT	-1.1256	17.2143	STDDEV	Y	
BMI	INPUT	11.1980	44.8472	STDDEV	Y	

SAS Enterprise Miner Report

**Node=Variable Selection
Summary**

Node id = Varsel
Node label = Variable Selection
Meta path = Ids => Filter => Varsel
Notes =

**Node=Variable Selection
Properties**

Property	Value	Default	Property	Value	Default	Property	Value	Default
Component	VariableSelection		MaxLevel	100		RoleUnusedVars	DEFAULT	
Bins	50		MaxMissingPercent	50		SASSPDS	Y	
ByPassVars	NONE		MaxRows	3000		StopR2	0.0005	
ByPassVarsRole	INPUT		MinR2	0.005		TargetModel	DEFAULT	
ChiSquare	3.84		Passes	6		UseAov16	N	
HideRejectedVars	Y		PrintOption	DEFAULT		UseGroups	Y	
HideUnusedInputVars	Y		RejectUnusedInputVars	Y		UseInteractions	N	

**Node=Variable Selection
Variable Summary**

Role	Level	Frequency Count	Name
INPUT	BINARY	14	AnyHealthcare CholCheck DiffWalk Fruits HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost PhysActivity Sex Smoker Stroke Veggies
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth

**Node=Variable Selection
Variable Selection**

Variable Name	Role	Measurement Level	Type	Label	Reasons for Rejection
Age	INPUT	INTERVAL	N		
AnyHealthcare	INPUT	BINARY	N		
BMI	INPUT	INTERVAL	N		
CholCheck	INPUT	BINARY	N		
DiffWalk	INPUT	BINARY	N		
Education	INPUT	ORDINAL	N		
Fruits	INPUT	BINARY	N		
GenHlth	INPUT	ORDINAL	N		
HeartDiseaseorAttack	INPUT	BINARY	N		
HighBP	INPUT	BINARY	N		
HighChol	INPUT	BINARY	N		
HvyAlcoholConsump	INPUT	BINARY	N		
Income	INPUT	ORDINAL	N		
MentHlth	INPUT	ORDINAL	N		
NoDocbcCost	INPUT	BINARY	N		
PhysActivity	INPUT	BINARY	N		
PhysHlth	INPUT	ORDINAL	N		

Variable Name	Role	Measurement Level	Type	Label	Reasons for Rejection
Sex	INPUT	BINARY	N		
Smoker	INPUT	BINARY	N		
Stroke	INPUT	BINARY	N		
Veggies	INPUT	BINARY	N		

SAS Enterprise Miner Report

Node=Average Clustering
Summary

Node id = Clus3
Node label = Average Clustering
Meta path = Ids => Filter => Varsel => Clus3
Notes =

Node=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	AVERAGE	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=Average Clustering
Variable Summary

Role	Level	Frequency Count	Name
INPUT	BINARY	14	AnyHealthcare CholCheck DiffWalk Fruits HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost PhysActivity Sex Smoker Stroke Veggies
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth

Node=Average Clustering
Mean Statistics

Variable	Segment 1	Segment 2	Segment 3	Segment 4
Clustering Criterion	0.70	0.70	0.70	0.70
Maximum Relative Change in Cluster Seeds	0.00	0.00	0.00	0.00
Improvement in Clustering Criterion	.	.	.	.
Frequency of Cluster	142732.00	14893.00	18895.00	71355.00
Root-Mean-Square Standard Deviation	0.61	0.90	0.83	0.78
Maximum Distance from Cluster Seed	7.86	9.91	9.43	9.19
Nearest Cluster	4.00	1.00	4.00	1.00
Distance to Nearest Cluster	3.13	4.14	4.01	3.13
Age	7.51	6.37	6.90	9.79
BMI	26.89	27.11	28.83	29.58
AnyHealthcare=0	0.00	0.49	0.20	0.01
AnyHealthcare=1	1.00	0.51	0.80	0.99
CholCheck=0	0.00	0.62	0.01	0.00

Variable	Segment 1	Segment 2	Segment 3	Segment 4
CholCheck=1	1.00	0.38	0.99	1.00
DiffWalk=0	0.99	0.94	0.68	0.57
DiffWalk=1	0.01	0.06	0.32	0.43
Fruits=0	0.31	0.41	0.43	0.44
Fruits=1	0.69	0.59	0.57	0.56
HeartDiseaseorAttack=0	0.98	0.98	0.87	0.75
HeartDiseaseorAttack=1	0.02	0.02	0.13	0.25
HighBP=0	0.72	0.80	0.53	0.26
HighBP=1	0.28	0.20	0.47	0.74
HighChol=0	0.67	0.78	0.54	0.35
HighChol=1	0.33	0.22	0.46	0.65
HvyAlcoholConsump=0	0.94	0.92	0.94	0.96
HvyAlcoholConsump=1	0.06	0.08	0.06	0.04
NoDocbcCost=0	1.00	0.90	-0.00	1.00
NoDocbcCost=1	0.00	0.10	1.00	0.00
PhysActivity=0	0.12	0.24	0.33	0.45
PhysActivity=1	0.88	0.76	0.67	0.55
Sex=0	0.55	0.49	0.63	0.57
Sex=1	0.45	0.51	0.37	0.43
Smoker=0	0.65	0.55	0.48	0.40
Smoker=1	0.35	0.45	0.52	0.60
Stroke=0	0.99	0.99	0.93	0.89
Stroke=1	0.01	0.01	0.07	0.11
Veggies=0	0.13	0.20	0.23	0.29
Veggies=1	0.87	0.80	0.77	0.71
Transform: Education	0.59	0.45	0.41	0.36
Transform: GenHlth	0.37	0.44	0.65	0.72
Transform: Income	0.62	0.44	0.31	0.33
Transform: MentHlth	0.46	0.50	0.64	0.55
Transform: PhysHlth	0.41	0.44	0.63	0.65
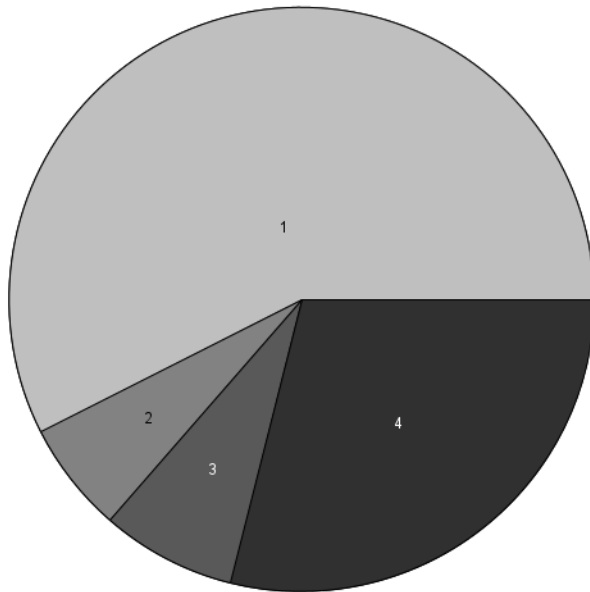
9

SAS Enterprise Miner Report
Node=Average Clustering
Segment Size

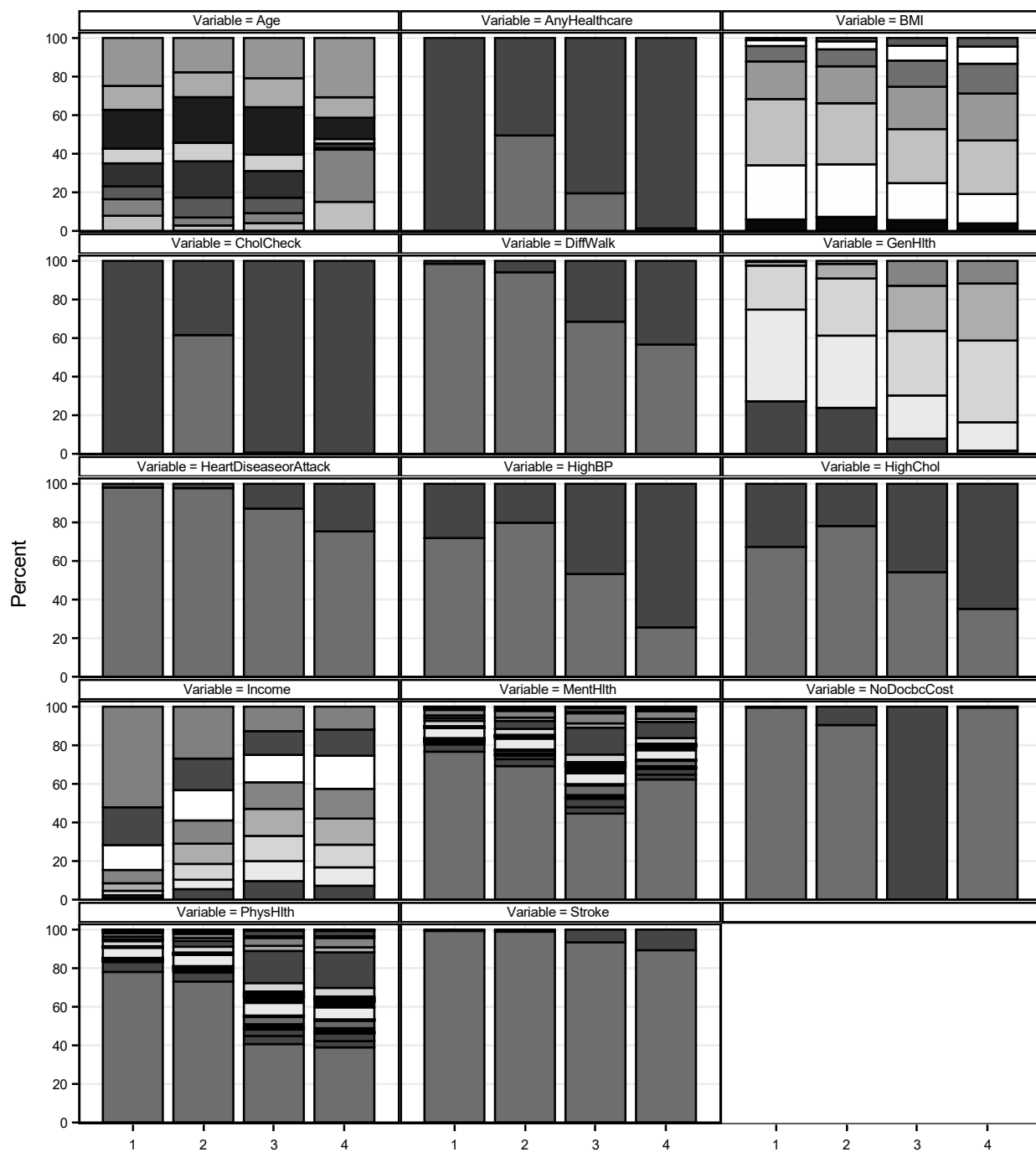SAS Enterprise Miner Report
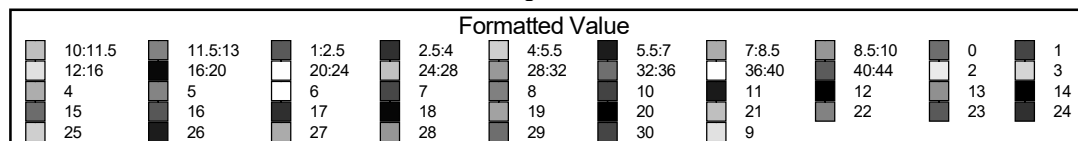Node=Average Clustering
Tree Diagram
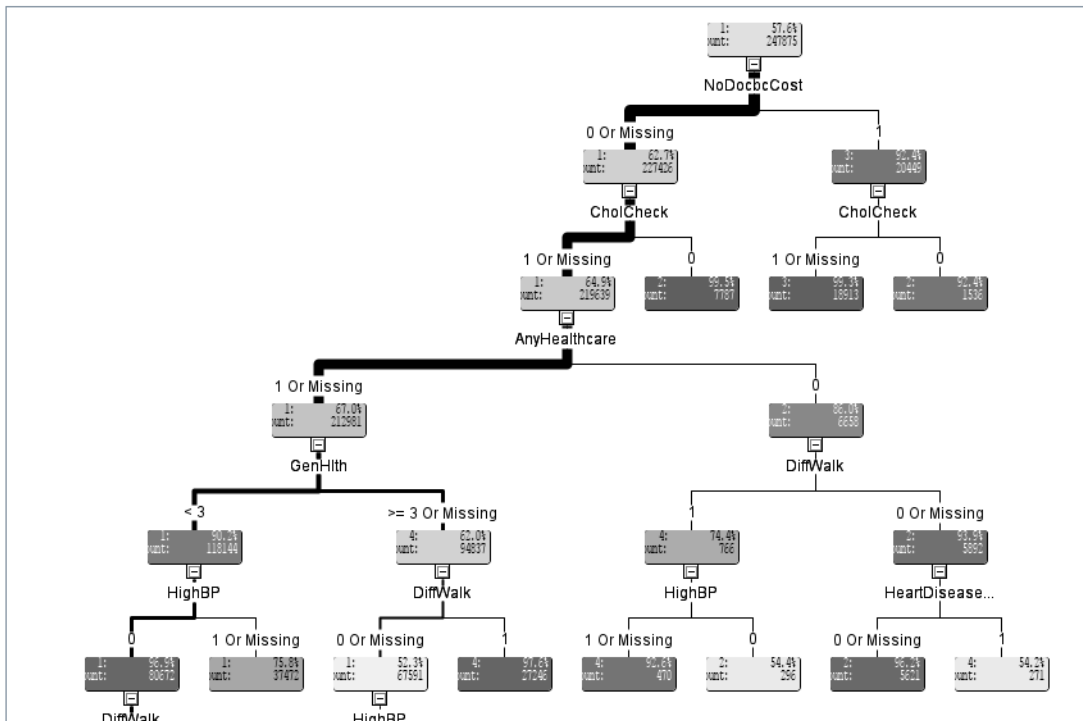
SAS Enterprise Miner Report
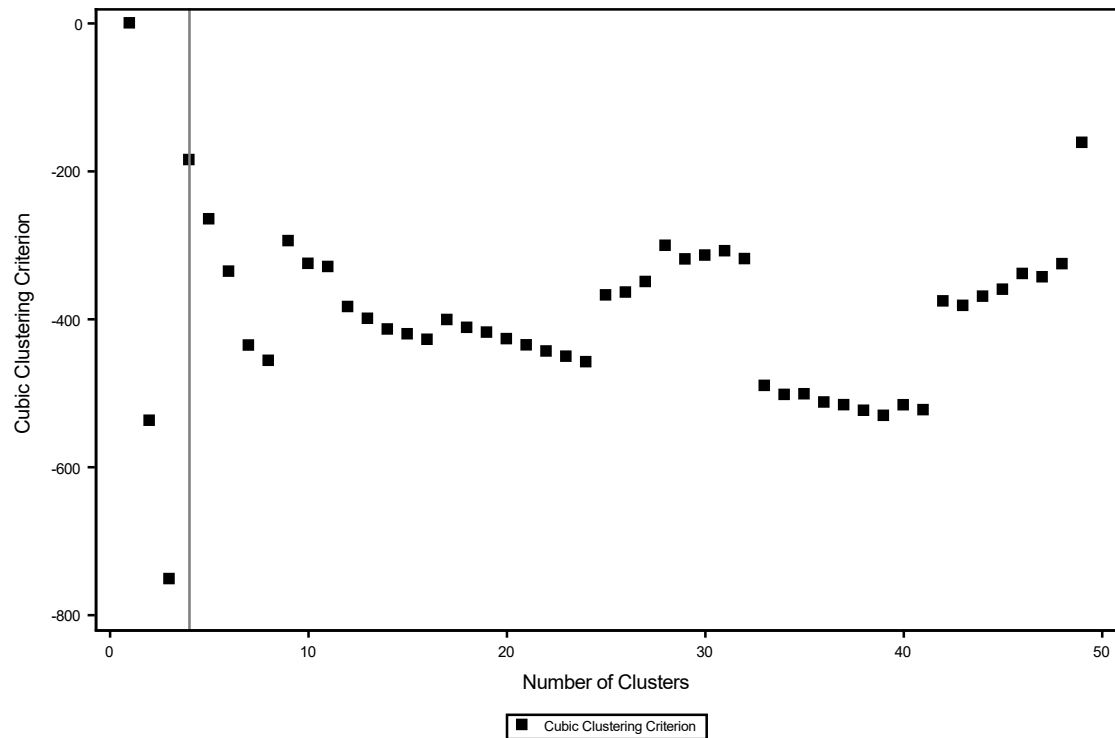Node=Average Clustering
Segment Plot

SAS Enterprise Miner Report
Node=Average Clustering
CCC Plot

SAS Enterprise Miner Report

Node=Centroid Clustering Summary

Node id = Clus2
Node label = Centroid Clustering
Meta path = Ids => Filter => Varsel => Clus2
Notes =

Node=Centroid 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=Centroid Clustering Variable Summary

Role	Level	Frequency Count	Name
INPUT	BINARY	14	AnyHealthcare CholCheck DiffWalk Fruits HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost PhysActivity Sex Smoker Stroke Veggies
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth

Node=Centroid Clustering Mean Statistics

Variable	Segment 1	Segment 2
Clustering Criterion	0.73	0.73
Maximum Relative Change in Cluster Seeds	0.00	0.00
Improvement in Clustering Criterion	.	.
Frequency of Cluster	79881.00	167994.00
Root-Mean-Square Standard Deviation	0.85	0.67
Maximum Distance from Cluster Seed	10.55	9.60
Nearest Cluster	2.00	1.00
Distance to Nearest Cluster	3.21	3.21
Age	9.34	7.44
BMI	29.70	26.94
AnyHealthcare=0	0.07	0.04
AnyHealthcare=1	0.93	0.96
CholCheck=0	0.02	0.05

Variable	Segment 1	Segment 2
CholCheck=1	0.98	0.95
DiffWalk=0	0.53	0.98
DiffWalk=1	0.47	0.02
Fruits=0	0.46	0.32
Fruits=1	0.54	0.68
HeartDiseaseorAttack=0	0.76	0.98
HeartDiseaseorAttack=1	0.24	0.02
HighBP=0	0.29	0.71
HighBP=1	0.71	0.29
HighChol=0	0.37	0.68
HighChol=1	0.63	0.32
HvyAlcoholConsump=0	0.96	0.94
HvyAlcoholConsump=1	0.04	0.06
NoDocbcCost=0	0.84	0.96
NoDocbcCost=1	0.16	0.04
PhysActivity=0	0.47	0.13
PhysActivity=1	0.53	0.87
Sex=0	0.59	0.54
Sex=1	0.41	0.46
Smoker=0	0.40	0.63
Smoker=1	0.60	0.37
Stroke=0	0.89	0.99
Stroke=1	0.11	0.01
Veggies=0	0.30	0.13
Veggies=1	0.70	0.87
Transform: Education	0.35	0.57
Transform: GenHlth	0.74	0.38
Transform: Income	0.30	0.59
Transform: MentHlth	0.58	0.46
Transform: PhysHlth	0.68	0.42

SAS Enterprise Miner Report
Node=Centroid Clustering
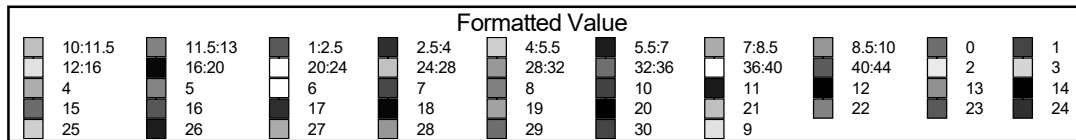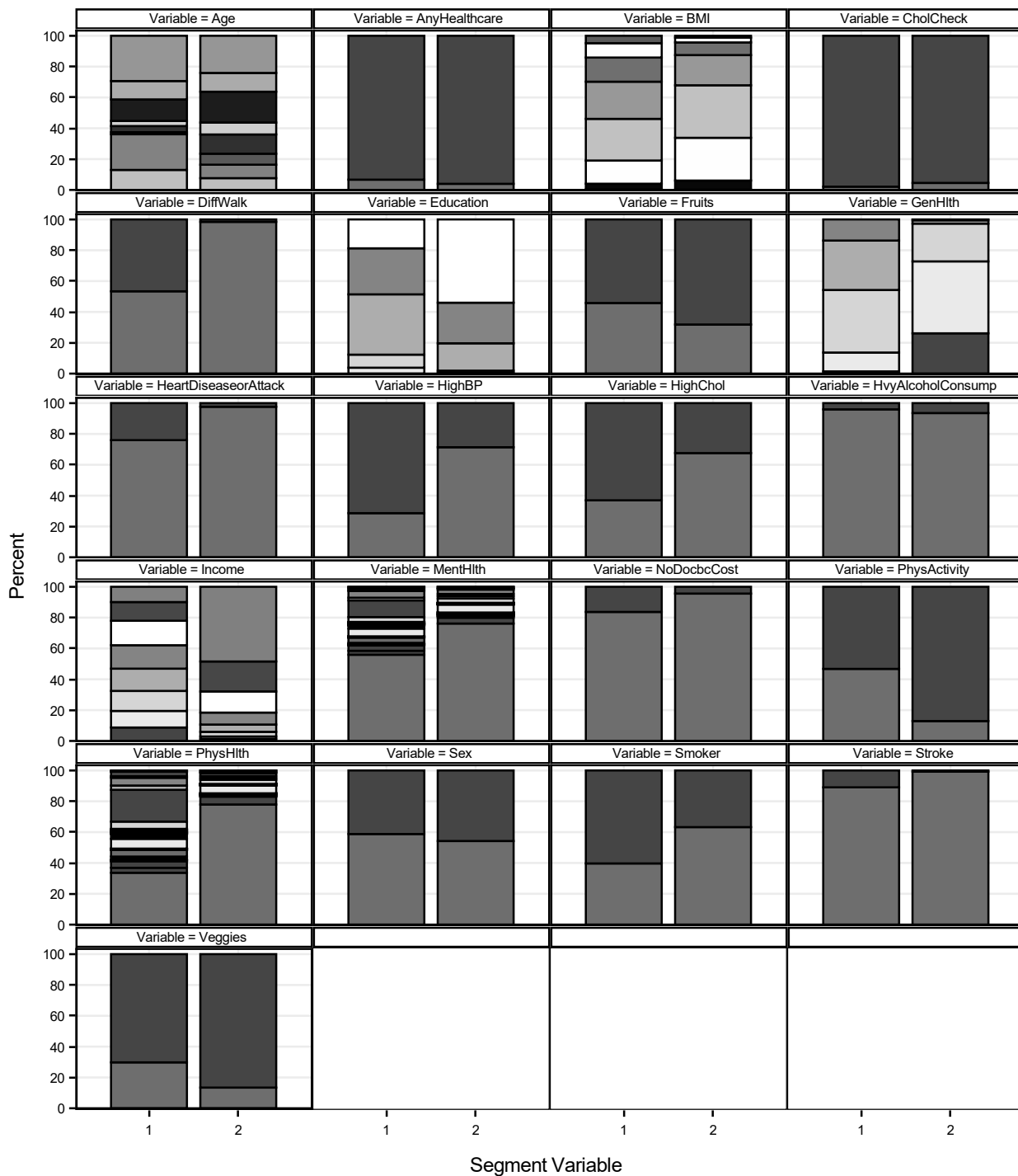### Segment Size
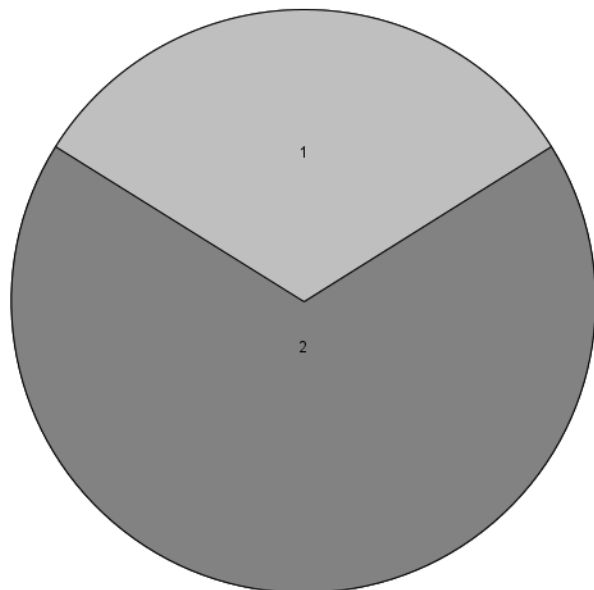
SAS Enterprise Miner Report
Node=Centroid Clustering
Tree Diagram
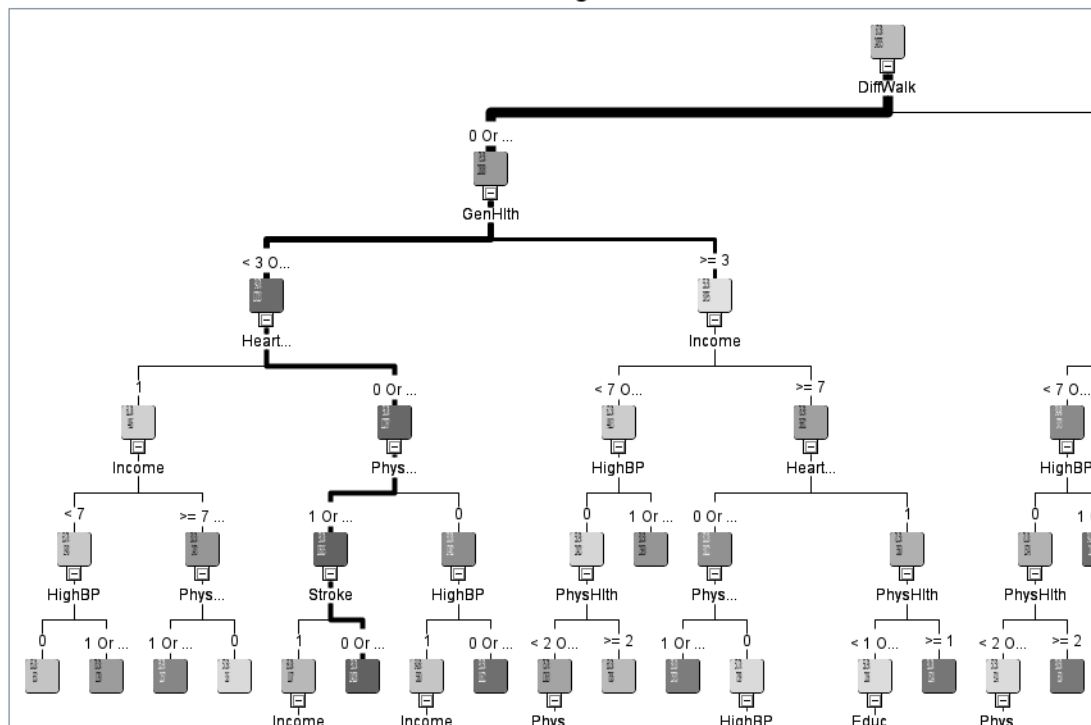
SAS Enterprise Miner Report
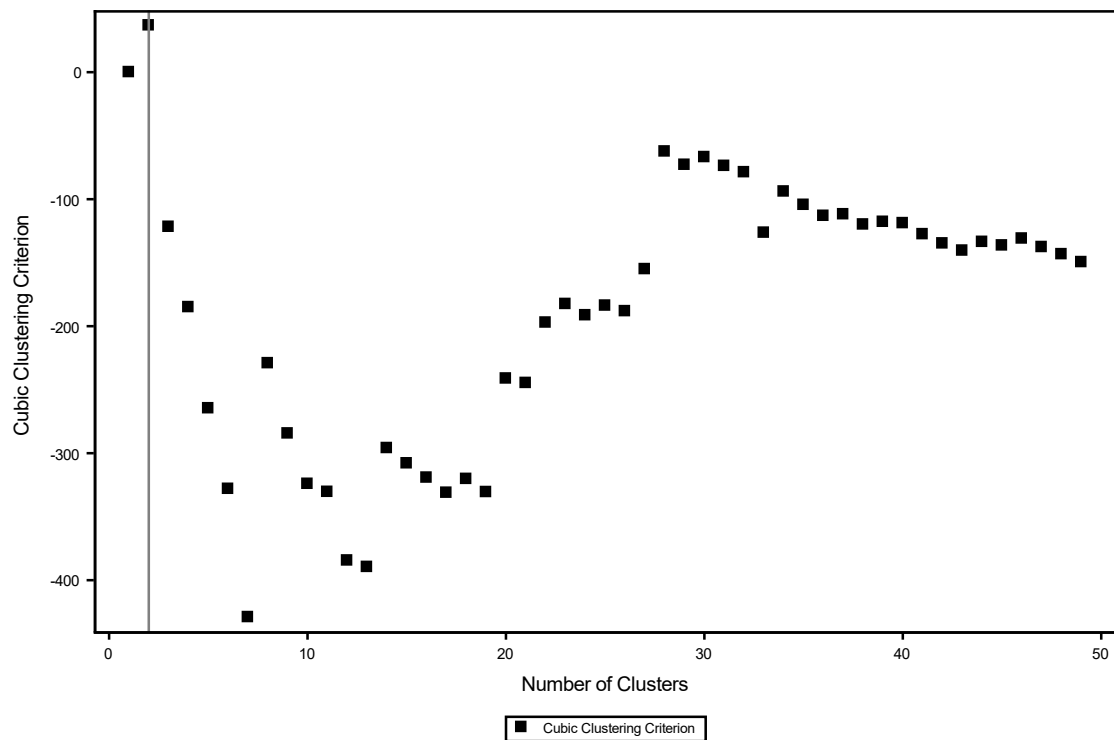## Node=Centroid Clustering
Segment Plot

SAS Enterprise Miner Report
Node=Centroid Clustering
CCC Plot

SAS Enterprise Miner Report

Node=Ward Clustering
Summary

Node id = Clus
Node label = Ward Clustering
Meta path = Ids => Filter => Varsel => Clus
Notes =

Node=Ward Clustering
Properties

Property	Value	Default	Property	Value	Default	Property	Value	Default
Component	Cluster		FinalMaxNum	20		MissingNominal	DEFAULT	
AutomaticMaxNum	50		HideVariable	Y		MissingOrdinal	DEFAULT	
AutomaticMethod	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=Ward Clustering
Variable Summary

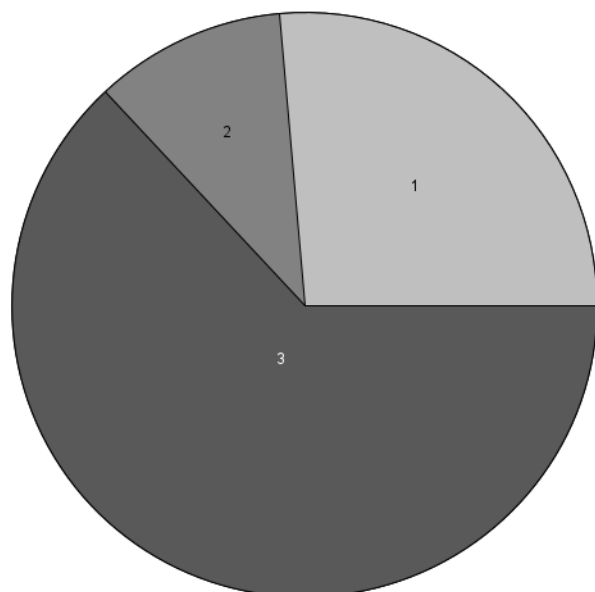| Role	Level	Frequency Count	Name
INPUT	BINARY	14	AnyHealthcare CholCheck DiffWalk Fruits HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost PhysActivity Sex Smoker Stroke Veggies
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth

Node=Ward Clustering
Mean Statistics

Variable	Segment 1	Segment 2	Segment 3
Clustering Criterion	0.71	0.71	0.71
Maximum Relative Change in Cluster Seeds	0.01	0.01	0.01
Improvement in Clustering Criterion	.	.	.
Frequency of Cluster	65173.00	26449.00	156253.00
Root-Mean-Square Standard Deviation	0.81	0.94	0.62
Maximum Distance from Cluster Seed	9.57	10.40	8.08
Nearest Cluster	3.00	3.00	1.00
Distance to Nearest Cluster	3.38	3.45	3.38
Age	9.70	6.15	7.69
BMI	29.77	27.58	27.05
AnyHealthcare=0	0.01	0.43	0.00
AnyHealthcare=1	0.99	0.57	1.00
CholCheck=0	0.00	0.34	0.00

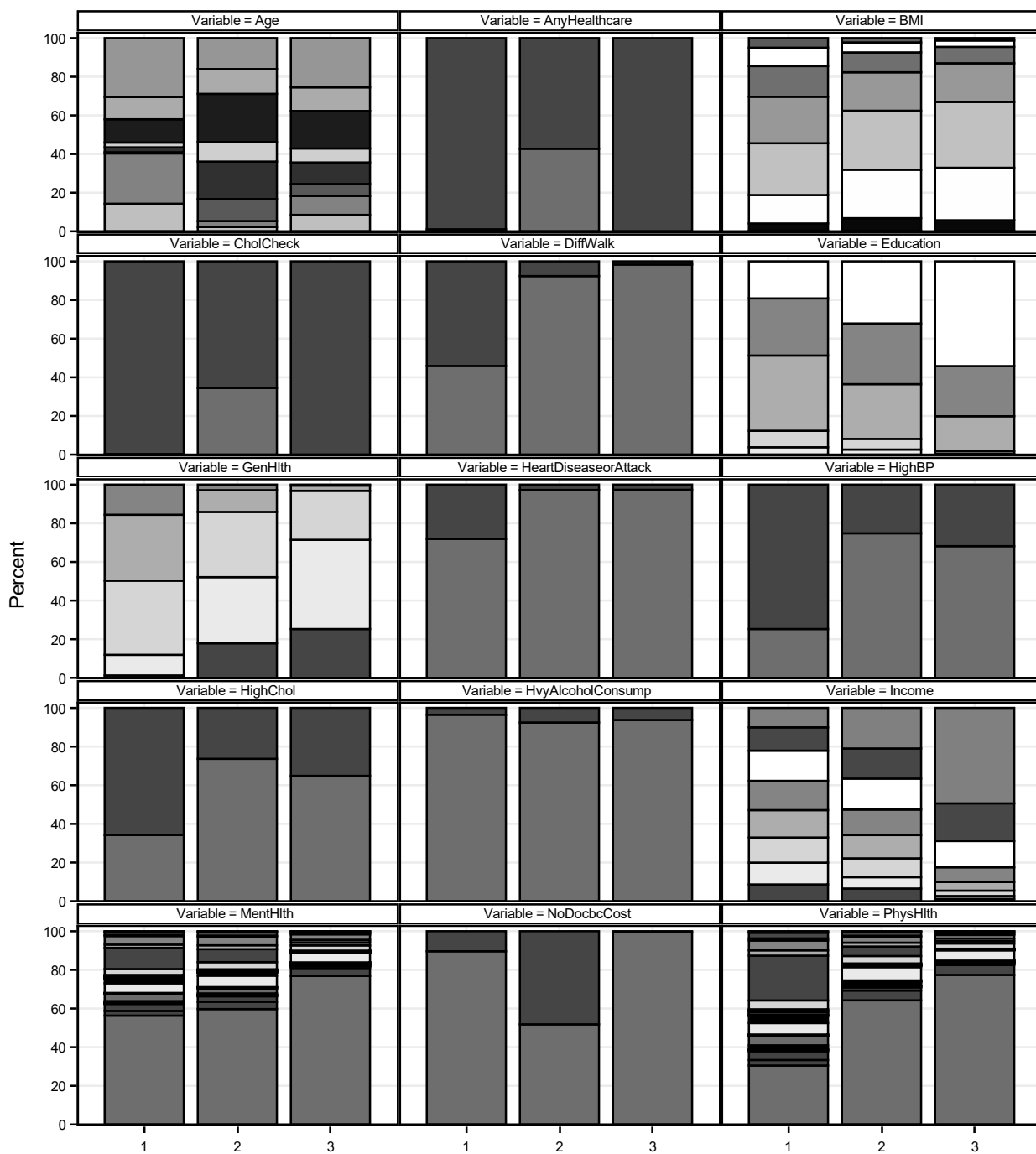| Variable	Segment 1	Segment 2	Segment 3
CholCheck=1	1.00	0.66	1.00
DiffWalk=0	0.46	0.92	0.98
DiffWalk=1	0.54	0.08	0.02
Fruits=0	0.45	0.41	0.32
Fruits=1	0.55	0.59	0.68
HeartDiseaseorAttack=0	0.72	0.97	0.97
HeartDiseaseorAttack=1	0.28	0.03	0.03
HighBP=0	0.25	0.75	0.68
HighBP=1	0.75	0.25	0.32
HighChol=0	0.34	0.74	0.65
HighChol=1	0.66	0.26	0.35
HvyAlcoholConsump=0	0.96	0.92	0.94
HvyAlcoholConsump=1	0.04	0.08	0.06
NoDocbcCost=0	0.90	0.52	0.99
NoDocbcCost=1	0.10	0.48	0.01
PhysActivity=0	0.49	0.24	0.13
PhysActivity=1	0.51	0.76	0.87
Sex=0	0.59	0.55	0.54
Sex=1	0.41	0.45	0.46
Smoker=0	0.39	0.54	0.63
Smoker=1	0.61	0.46	0.37
Stroke=0	0.87	0.99	0.99
Stroke=1	0.13	0.01	0.01
Veggies=0	0.30	0.20	0.14
Veggies=1	0.70	0.80	0.86
Transform: Education	0.35	0.44	0.57
Transform: GenHlth	0.76	0.50	0.39
Transform: Income	0.30	0.40	0.60
Transform: MentHlth	0.58	0.55	0.46
Transform: PhysHlth	0.70	0.49	0.42

SAS Enterprise Miner Report
Node=Ward Clustering
Segment Size

SAS Enterprise Miner Report
Node=Ward Clustering
Tree Diagram

SAS Enterprise Miner Report
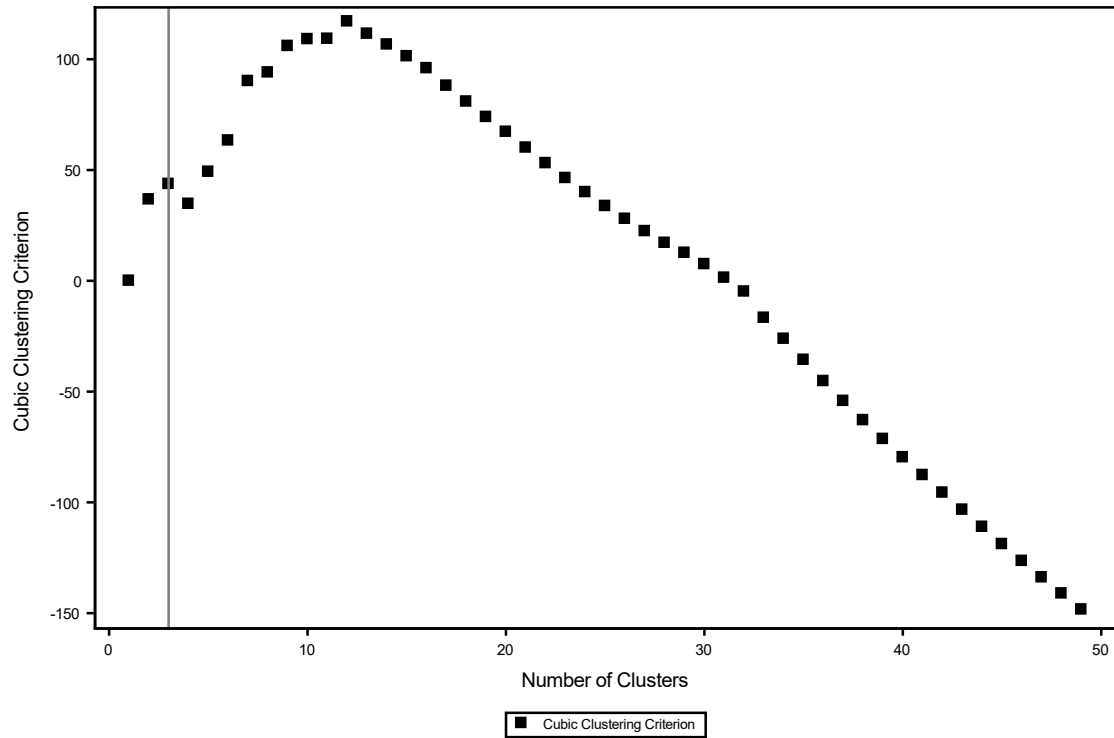Node=Ward Clustering
Segment Plot

SAS Enterprise Miner Report
Node=Ward Clustering
CCC Plot

SAS Enterprise Miner Report

Node=Sample(90%)
Summary

Node id = Smpl
Node label = Sample(90%)
Meta path = Ids => 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	90	10
FreqCount	N		Method	DEFAULT		SizeType	PERCENT	
FreqMiss	N		MinStrataSize	5		StratifyCriterion	PROPORTIONAL	

Node=Sample(90%)
Variable Summary

Role	Level	Frequency Count	Name
REJECTED	NOMINAL	1	Diabetes_012
INPUT	BINARY	14	AnyHealthcare CholCheck DiffWalk Fruits HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost PhysActivity Sex Smoker Stroke Veggies
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth

SAS Enterprise Miner Report

Node=Variable Selection (2)
Summary

Node id = Varsel2
Node label = Variable Selection (2)
Meta path = lds => Filter => Smpl => Varsel2
Notes =

Node=Variable Selection (2)
Properties

Property	Value	Default	Property	Value	Default	Property	Value	Default
Component	VariableSelection		MaxLevel	100		RoleUnusedVars	DEFAULT	
Bins	50		MaxMissingPercent	50		SASSPDS	Y	
ByPassVars	NONE		MaxRows	3000		StopR2	0.0005	
ByPassVarsRole	INPUT		MinR2	0.005		TargetModel	DEFAULT	
ChiSquare	3.84		Passes	6		UseAov16	N	
HideRejectedVars	Y		PrintOption	DEFAULT		UseGroups	Y	
HideUnusedInputVars	Y		RejectUnusedInputVars	Y		UseInteractions	N	

Node=Variable Selection (2)
Variable Summary

Role	Level	Frequency Count	Name
INPUT	BINARY	14	AnyHealthcare CholCheck DiffWalk Fruits HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost PhysActivity Sex Smoker Stroke Veggies
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth

Node=Variable Selection (2)
Variable Selection

Variable Name	Role	Measurement Level	Type	Label	Reasons for Rejection
Age	INPUT	INTERVAL	N		
AnyHealthcare	INPUT	BINARY	N		
BMI	INPUT	INTERVAL	N		
CholCheck	INPUT	BINARY	N		
DiffWalk	INPUT	BINARY	N		
Education	INPUT	ORDINAL	N		
Fruits	INPUT	BINARY	N		
GenHlth	INPUT	ORDINAL	N		
HeartDiseaseorAttack	INPUT	BINARY	N		
HighBP	INPUT	BINARY	N		
HighChol	INPUT	BINARY	N		
HvyAlcoholConsump	INPUT	BINARY	N		
Income	INPUT	ORDINAL	N		
MentHlth	INPUT	ORDINAL	N		
NoDocbcCost	INPUT	BINARY	N		
PhysActivity	INPUT	BINARY	N		
PhysHlth	INPUT	ORDINAL	N		

Variable Name	Role	Measurement Level	Type	Label	Reasons for Rejection
Sex	INPUT	BINARY	N		
Smoker	INPUT	BINARY	N		
Stroke	INPUT	BINARY	N		
Veggies	INPUT	BINARY	N		

SAS Enterprise Miner Report

**Node=90% Average Clustering
Summary**

Node id = Clus6
Node label = 90% Average Clustering
Meta path = Ids => Filter => Smpl => Varsel2 => 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	AVERAGE	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	BINARY	14	AnyHealthcare CholCheck DiffWalk Fruits HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost PhysActivity Sex Smoker Stroke Veggies
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth
ID	INTERVAL	1	_dataobs_

**Node=90% Average Clustering
Mean Statistics**

Variable	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5
Clustering Criterion	0.68	0.68	0.68	0.68	0.68
Maximum Relative Change in Cluster Seeds	0.00	0.00	0.00	0.00	0.00
Improvement in Clustering Criterion	.	.	.	.	.
Frequency of Cluster	36091.00	8401.00	91277.00	11802.00	75517.00
Root-Mean-Square Standard Deviation	0.83	0.77	0.58	0.70	0.69
Maximum Distance from Cluster Seed	9.60	10.16	8.69	9.14	9.04
Nearest Cluster	5.00	3.00	5.00	5.00	3.00
Distance to Nearest Cluster	3.14	5.43	2.41	4.54	2.41
Age	9.39	6.65	6.65	7.64	9.32
BMI	29.97	26.79	26.20	26.76	29.05
AnyHealthcare=0	0.05	0.18	0.04	0.05	0.05
AnyHealthcare=1	0.95	0.82	0.96	0.95	0.95

Variable	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5
CholCheck=0	0.00	1.00	0.00	0.00	-0.00
CholCheck=1	1.00	0.00	1.00	1.00	1.00
DiffWalk=0	0.15	0.91	0.98	0.90	0.98
DiffWalk=1	0.85	0.09	0.02	0.10	0.02
Fruits=0	0.43	0.42	0.27	0.43	0.43
Fruits=1	0.57	0.58	0.73	0.57	0.57
HeartDiseaseorAttack=0	0.71	0.97	0.99	0.95	0.88
HeartDiseaseorAttack=1	0.29	0.03	0.01	0.05	0.12
HighBP=0	0.31	0.82	0.88	0.57	0.30
HighBP=1	0.69	0.18	0.12	0.43	0.70
HighChol=0	0.39	0.79	0.79	0.59	0.38
HighChol=1	0.61	0.21	0.21	0.41	0.62
HvyAlcoholConsump=0	1.00	0.92	1.00	0.00	1.00
HvyAlcoholConsump=1	0.00	0.08	-0.00	1.00	-0.00
NoDocbcCost=0	0.81	0.84	0.95	0.92	0.94
NoDocbcCost=1	0.19	0.16	0.05	0.08	0.06
PhysActivity=0	0.52	0.25	0.10	0.21	0.28
PhysActivity=1	0.48	0.75	0.90	0.79	0.72
Sex=0	0.64	0.50	0.58	0.55	0.49
Sex=1	0.36	0.50	0.42	0.45	0.51
Smoker=0	0.40	0.53	0.72	0.36	0.46
Smoker=1	0.60	0.47	0.28	0.64	0.54
Stroke=0	0.84	0.99	0.99	0.98	0.97
Stroke=1	0.16	0.01	0.01	0.02	0.03
Veggies=0	0.28	0.20	0.10	0.15	0.25
Veggies=1	0.72	0.80	0.90	0.85	0.75
Transform: Education	0.36	0.49	0.62	0.53	0.41
Transform: GenHlth	0.80	0.43	0.34	0.46	0.57
Transform: Income	0.28	0.48	0.64	0.57	0.43
Transform: MentHlth	0.64	0.52	0.48	0.53	0.46
Transform: PhysHlth	0.78	0.46	0.42	0.47	0.47

27

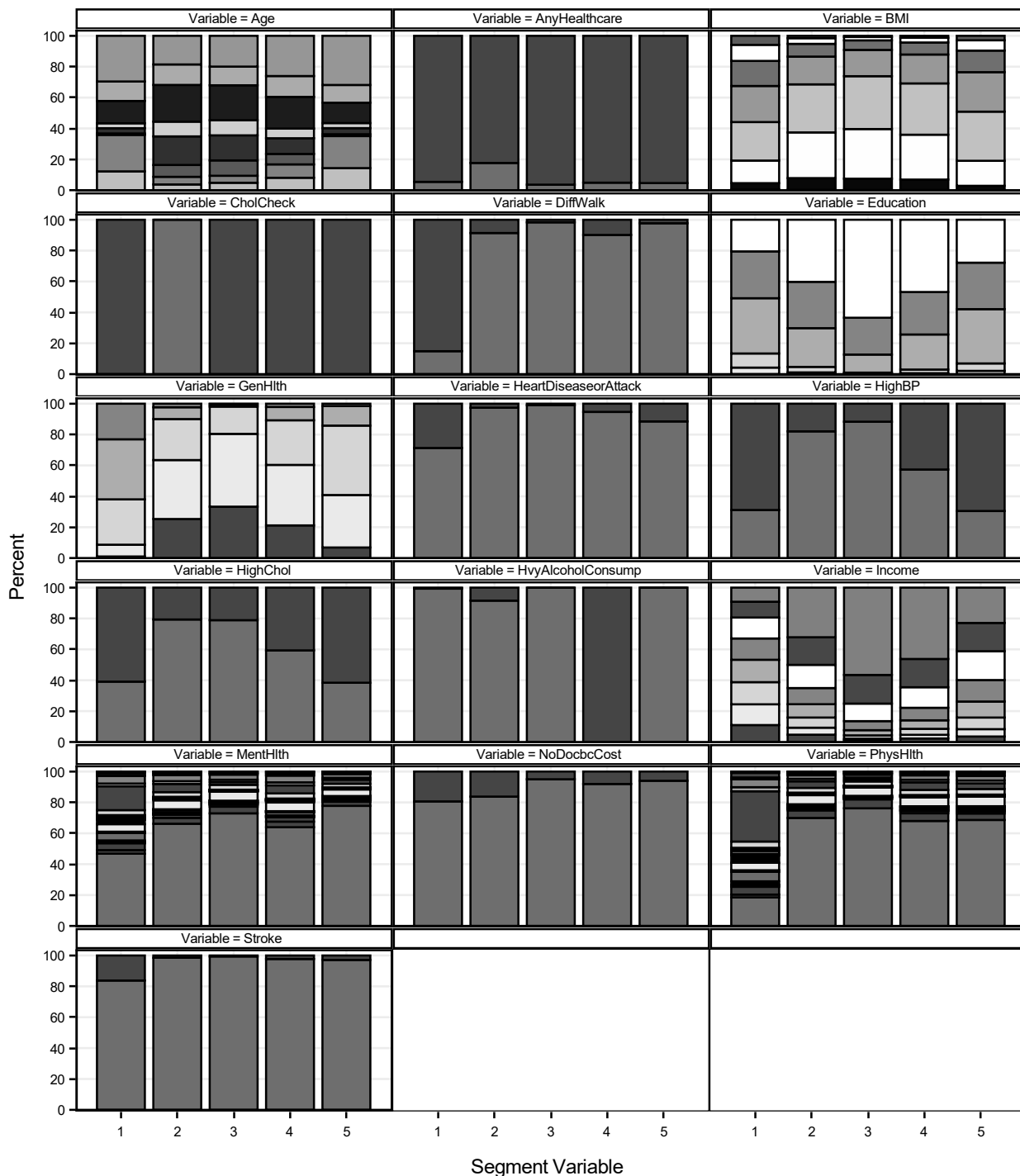SAS Enterprise Miner Report
Node=90% Average Clustering
Segment Size
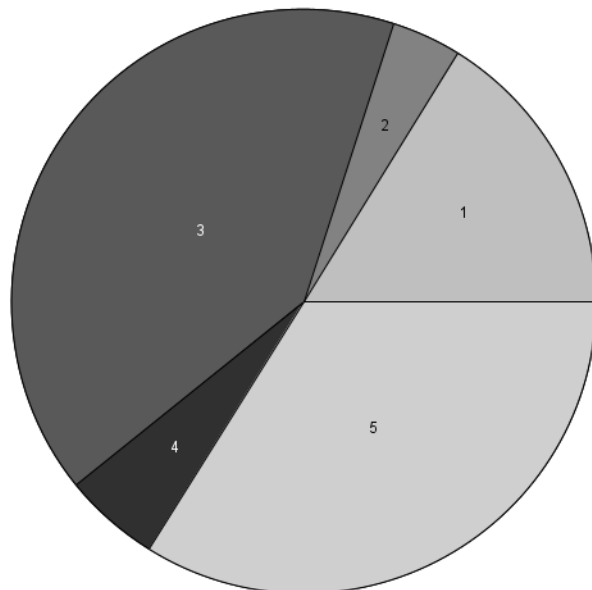
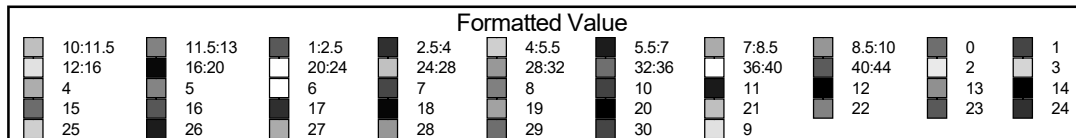SAS Enterprise Miner Report
Node=90% Average Clustering
Tree Diagram
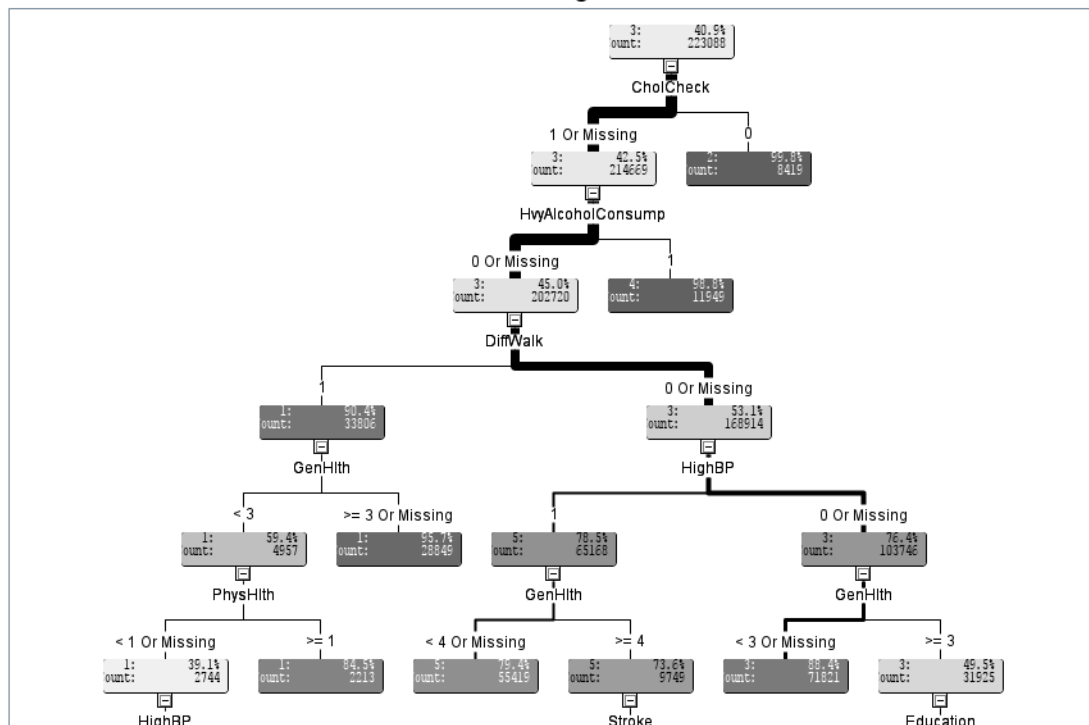
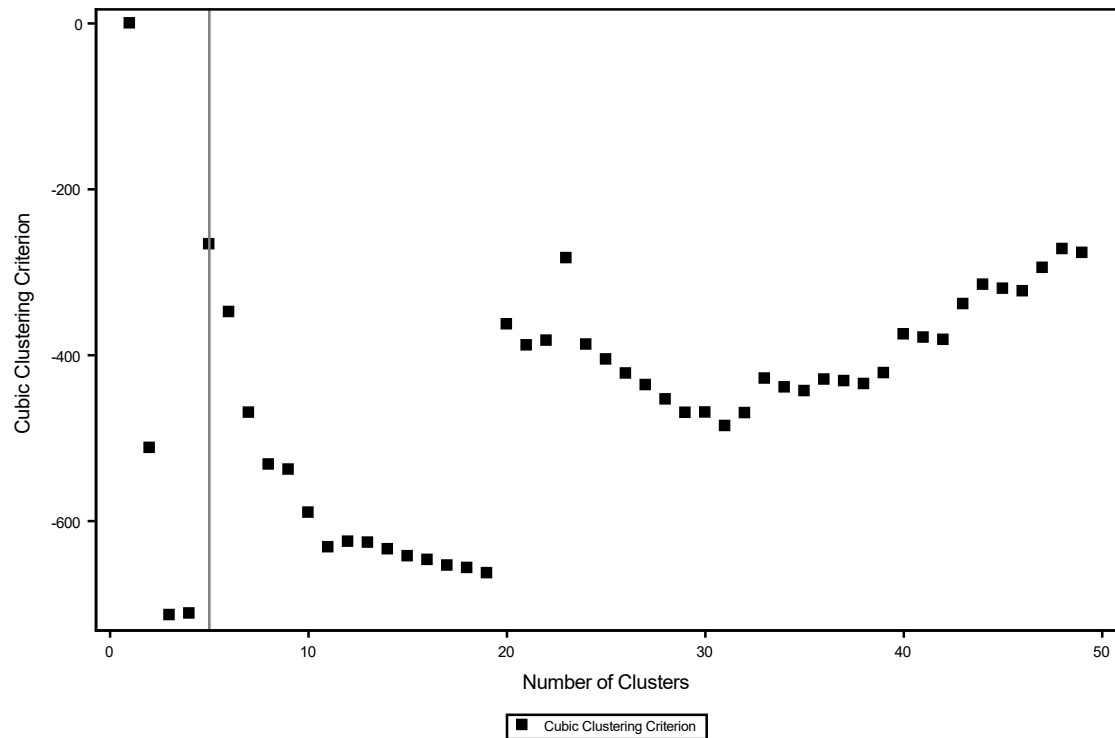SAS Enterprise Miner Report
Node=90% Average Clustering
Segment Plot

SAS Enterprise Miner Report
Node=90% Average Clustering
CCC Plot

SAS Enterprise Miner Report

**Node=90% Centroid Clustering
Summary**

Node id = Clus5
Node label = 90% Centroid Clustering
Meta path = Ids => Filter => Smpl => Varsel2 => Clus5
Notes =

**Node=90% Centroid 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% Centroid Clustering
Variable Summary**

Role	Level	Frequency Count	Name
INPUT	BINARY	14	AnyHealthcare CholCheck DiffWalk Fruits HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost PhysActivity Sex Smoker Stroke Veggies
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth
ID	INTERVAL	1	_dataobs_

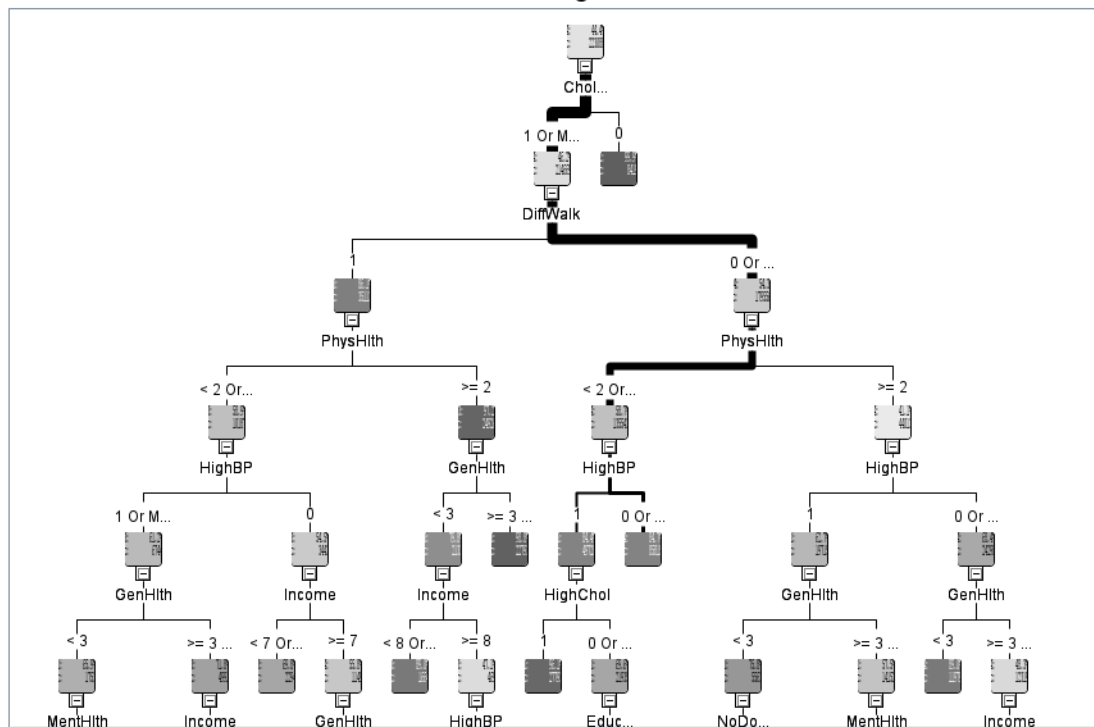**Node=90% Centroid Clustering
Mean Statistics**

Variable	Segment 1	Segment 2	Segment 3	Segment 4
Clustering Criterion	0.69	0.69	0.69	0.69
Maximum Relative Change in Cluster Seeds	0.01	0.01	0.01	0.01
Improvement in Clustering Criterion	.	.	.	.
Frequency of Cluster	73201.00	8408.00	42407.00	99072.00
Root-Mean-Square Standard Deviation	0.68	0.77	0.85	0.62
Maximum Distance from Cluster Seed	9.28	10.16	9.61	8.68
Nearest Cluster	4.00	4.00	1.00	1.00
Distance to Nearest Cluster	2.43	5.39	3.13	2.43
Age	9.74	6.65	8.97	6.52
BMI	28.79	26.79	29.94	26.30
AnyHealthcare=0	0.02	0.18	0.08	0.04
AnyHealthcare=1	0.98	0.82	0.92	0.96

Variable	Segment 1	Segment 2	Segment 3	Segment 4
CholCheck=0	-0.00	1.00	0.00	0.00
CholCheck=1	1.00	0.00	1.00	1.00
DiffWalk=0	0.95	0.91	0.29	0.98
DiffWalk=1	0.05	0.09	0.71	0.02
Fruits=0	0.38	0.42	0.46	0.30
Fruits=1	0.62	0.58	0.54	0.70
HeartDiseaseorAttack=0	0.87	0.97	0.75	0.99
HeartDiseaseorAttack=1	0.13	0.03	0.25	0.01
HighBP=0	0.22	0.82	0.35	0.92
HighBP=1	0.78	0.18	0.65	0.08
HighChol=0	0.34	0.79	0.42	0.80
HighChol=1	0.66	0.21	0.58	0.20
HvyAlcoholConsump=0	0.94	0.92	0.96	0.94
HvyAlcoholConsump=1	0.06	0.08	0.04	0.06
NoDocbcCost=0	0.97	0.84	0.77	0.95
NoDocbcCost=1	0.03	0.16	0.23	0.05
PhysActivity=0	0.24	0.25	0.52	0.12
PhysActivity=1	0.76	0.75	0.48	0.88
Sex=0	0.46	0.50	0.64	0.59
Sex=1	0.54	0.50	0.36	0.41
Smoker=0	0.47	0.53	0.39	0.69
Smoker=1	0.53	0.47	0.61	0.31
Stroke=0	0.96	0.98	0.86	1.00
Stroke=1	0.04	0.02	0.14	0.00
Veggies=0	0.21	0.20	0.30	0.12
Veggies=1	0.79	0.80	0.70	0.88
Transform: Education	0.45	0.49	0.35	0.60
Transform: GenHlth	0.54	0.43	0.80	0.35
Transform: Income	0.47	0.48	0.27	0.62
Transform: MentHlth	0.43	0.52	0.66	0.48
Transform: PhysHlth	0.44	0.46	0.79	0.43

SAS Enterprise Miner Report
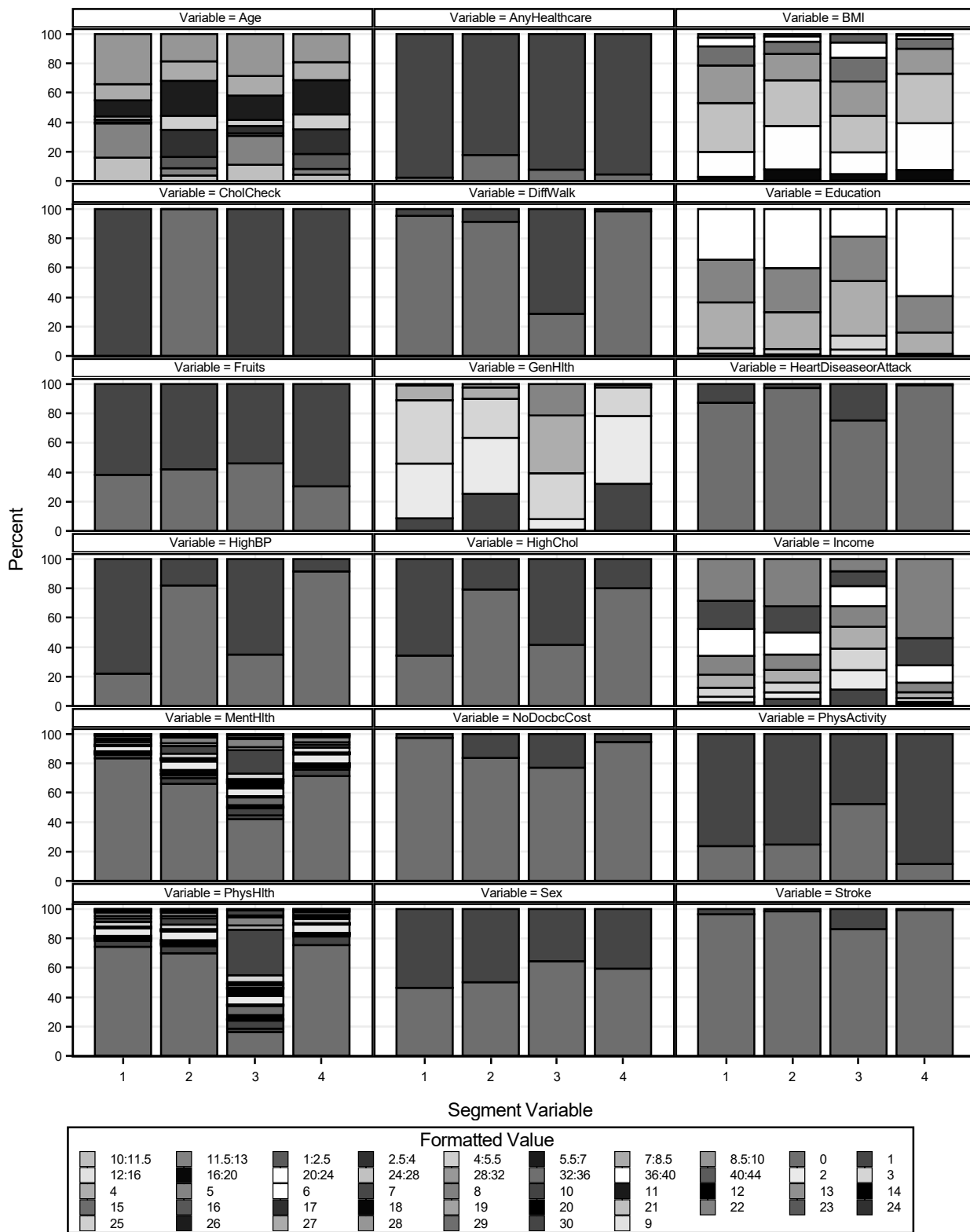Node=90% Centroid Clustering
Segment Size
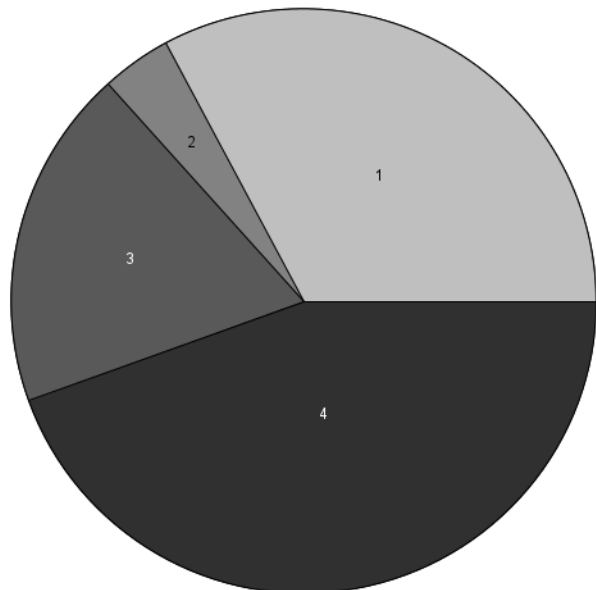
SAS Enterprise Miner Report
Node=90% Centroid Clustering
Tree Diagram

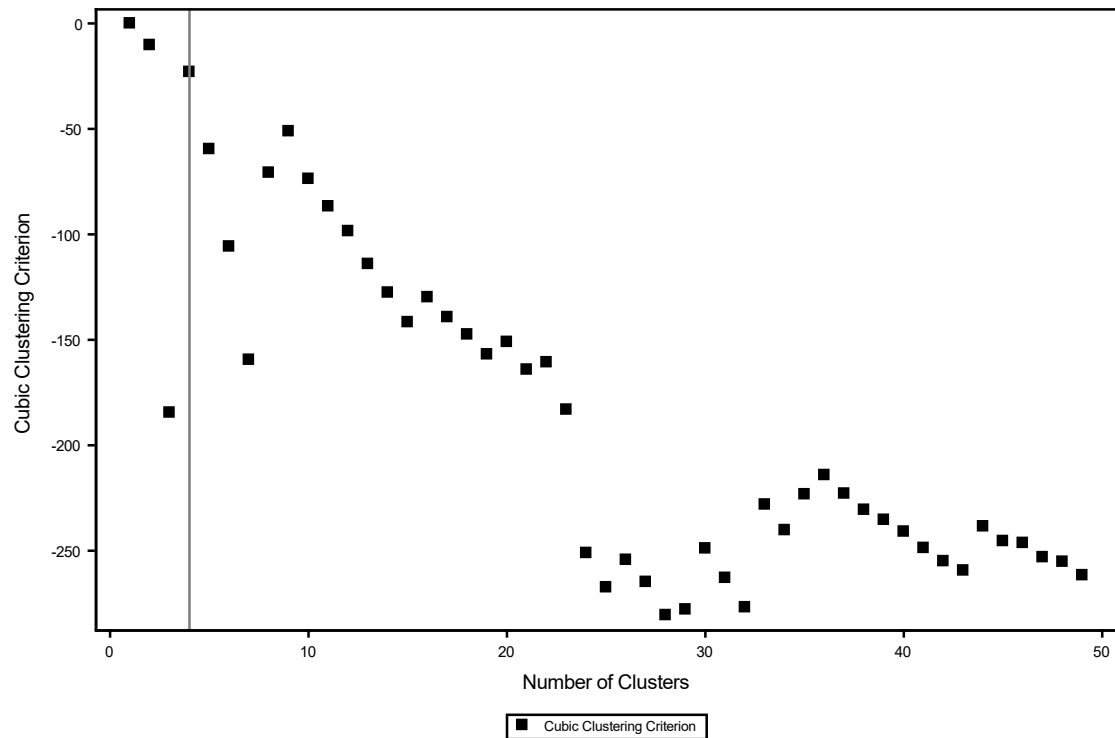# SAS Enterprise Miner Report
Node=90% Centroid Clustering
Segment Plot

SAS Enterprise Miner Report
Node=90% Centroid Clustering
CCC Plot

SAS Enterprise Miner Report

Node=90% Ward Clustering
Summary

Node id = Clus4
Node label = 90% Ward Clustering
Meta path = lds => Filter => Smpl => Varsel2 => Clus4
Notes =

Node=90% Ward Clustering
Properties

Property	Value	Default	Property	Value	Default	Property	Value	Default
Component	Cluster		FinalMaxNum	20		MissingNominal	DEFAULT	
AutomaticMaxNum	50		HideVariable	Y		MissingOrdinal	DEFAULT	
AutomaticMethod	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% Ward Clustering
Variable Summary

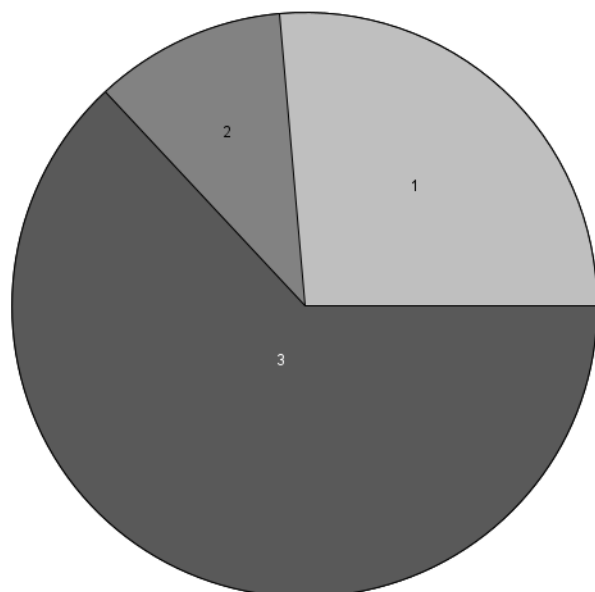| Role	Level	Frequency Count	Name
INPUT	BINARY	14	AnyHealthcare CholCheck DiffWalk Fruits HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost PhysActivity Sex Smoker Stroke Veggies
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth
ID	INTERVAL	1	_dataobs_

Node=90% Ward Clustering
Mean Statistics

Variable	Segment 1	Segment 2	Segment 3
Clustering Criterion	0.71	0.71	0.71
Maximum Relative Change in Cluster Seeds	0.01	0.01	0.01
Improvement in Clustering Criterion	.	.	.
Frequency of Cluster	58849.00	23775.00	140464.00
Root-Mean-Square Standard Deviation	0.81	0.94	0.62
Maximum Distance from Cluster Seed	9.58	10.40	7.86
Nearest Cluster	3.00	3.00	1.00
Distance to Nearest Cluster	3.37	3.45	3.37
Age	9.69	6.14	7.69
BMI	29.78	27.56	27.06
AnyHealthcare=0	0.01	0.43	0.00
AnyHealthcare=1	0.99	0.57	1.00

Variable	Segment 1	Segment 2	Segment 3
CholCheck=0	0.00	0.35	-0.00
CholCheck=1	1.00	0.65	1.00
DiffWalk=0	0.46	0.92	0.98
DiffWalk=1	0.54	0.08	0.02
Fruits=0	0.45	0.41	0.32
Fruits=1	0.55	0.59	0.68
HeartDiseaseorAttack=0	0.72	0.97	0.97
HeartDiseaseorAttack=1	0.28	0.03	0.03
HighBP=0	0.25	0.75	0.68
HighBP=1	0.75	0.25	0.32
HighChol=0	0.34	0.74	0.65
HighChol=1	0.66	0.26	0.35
HvyAlcoholConsump=0	0.96	0.92	0.94
HvyAlcoholConsump=1	0.04	0.08	0.06
NoDocbcCost=0	0.90	0.52	0.99
NoDocbcCost=1	0.10	0.48	0.01
PhysActivity=0	0.48	0.24	0.13
PhysActivity=1	0.52	0.76	0.87
Sex=0	0.59	0.55	0.54
Sex=1	0.41	0.45	0.46
Smoker=0	0.39	0.54	0.63
Smoker=1	0.61	0.46	0.37
Stroke=0	0.87	0.99	0.99
Stroke=1	0.13	0.01	0.01
Veggies=0	0.30	0.20	0.14
Veggies=1	0.70	0.80	0.86
Transform: Education	0.35	0.44	0.57
Transform: GenHlth	0.76	0.50	0.39
Transform: Income	0.30	0.40	0.60
Transform: MentHlth	0.58	0.55	0.46
Transform: PhysHlth	0.70	0.49	0.42

	Segment 1	Segment 2	Segment 3

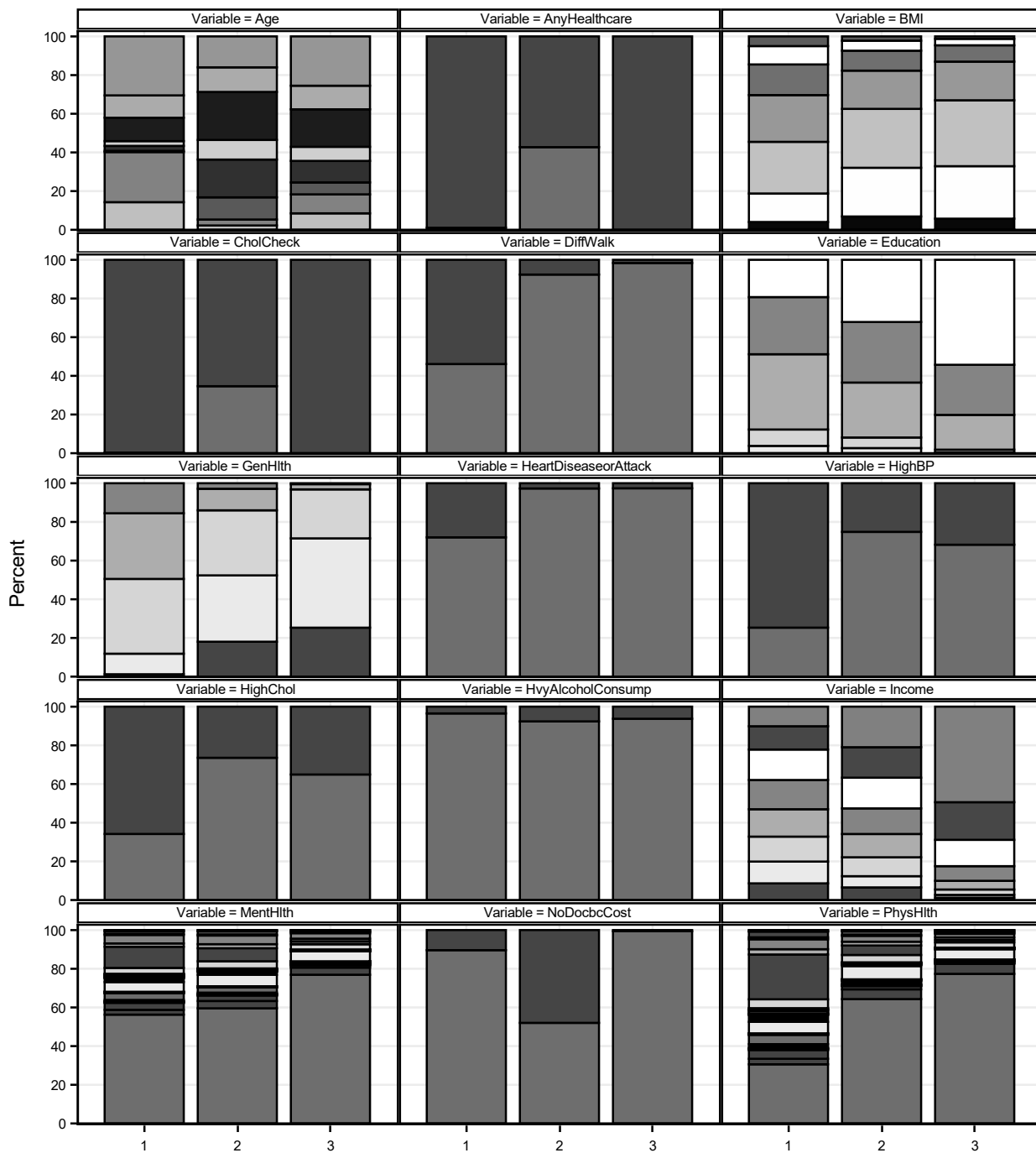SAS Enterprise Miner Report
Node=90% Ward Clustering
Segment Size

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

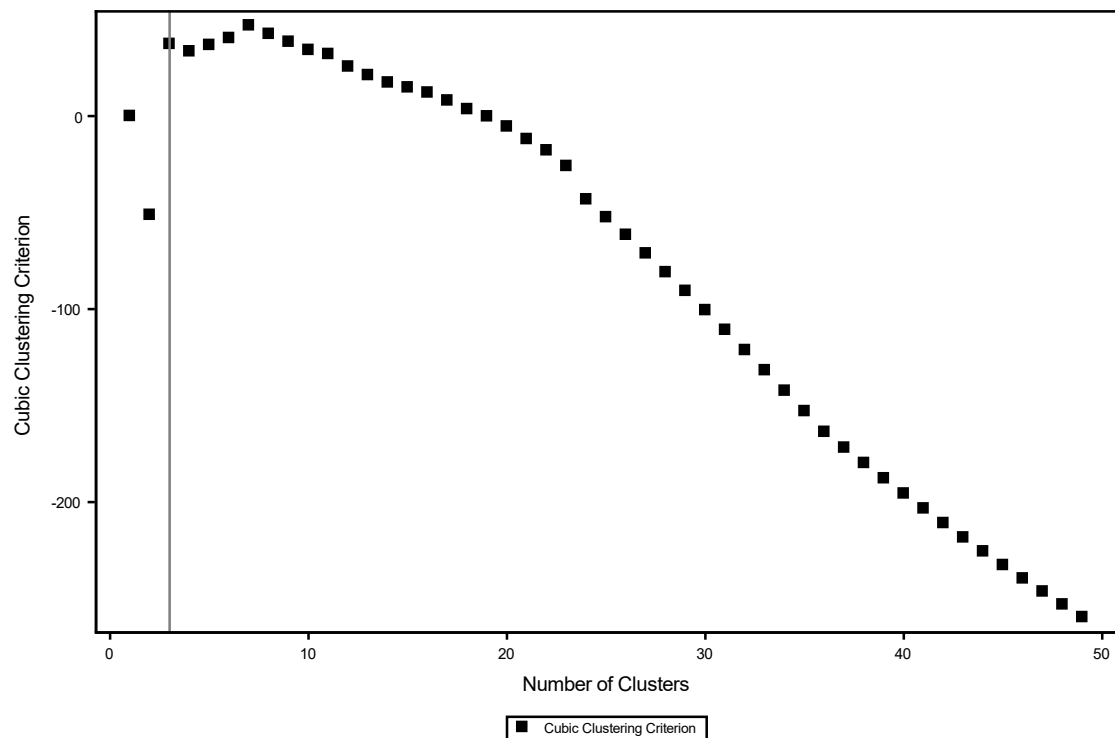SAS Enterprise Miner Report
Node=90% Ward Clustering
Segment Plot

SAS Enterprise Miner Report
Node=90% Ward Clustering
CCC Plot

End of Report