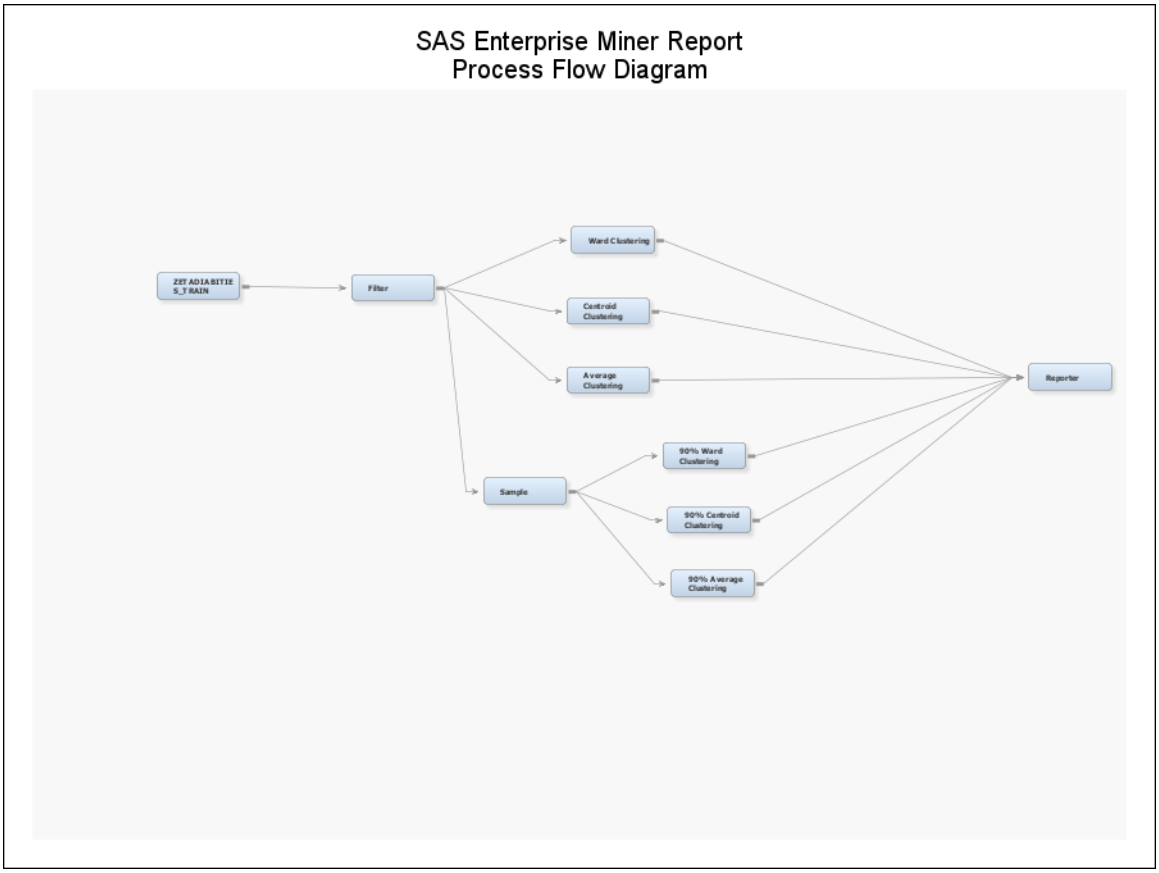
SAS Enterprise Miner Report

User = bkris2
Date = 08:57:46 March 04
Project = ZetaExploratoryAnalysis
Diagram = ZetaExploratoryAnalysis

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

Format = PDF
Style = LISTING



SAS Enterprise Miner Report

Node=ZETADIABITIES_TRAIN Summary

Node id = Ids
Node label = ZETADIABITIES_TRAIN
Meta path = Ids
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		REJECTED	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		REJECTED	BINARY	N	8	BEST12.0	
PhysHlth		INPUT	ORDINAL	N	8	BEST12.0	
Sex		REJECTED	BINARY	N	8	BEST12.0	
Smoker		REJECTED	BINARY	N	8	BEST12.0	
Stroke		REJECTED	BINARY	N	8	BEST12.0	
Veggies		REJECTED	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	BINARY	6	Fruits PhysActivity Sex Smoker Stroke Veggies
REJECTED	NOMINAL	1	Diabetes_012
INPUT	BINARY	8	AnyHealthcare CholCheck DiffWalk HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost
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=Sample
Summary

Node id = Smpl
Node label = Sample
Meta path = Ids => Filter => Smpl
Notes =

Node=Sample
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
Variable Summary

Role	Level	Frequency Count	Name
REJECTED	BINARY	6	Fruits PhysActivity Sex Smoker Stroke Veggies
REJECTED	NOMINAL	1	Diabetes_012
INPUT	BINARY	8	AnyHealthcare CholCheck DiffWalk HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth

SAS Enterprise Miner Report

Node=90% Average Clustering Summary

Node id = Clus6
Node label = 90% Average Clustering
Meta path = Ids => 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	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				

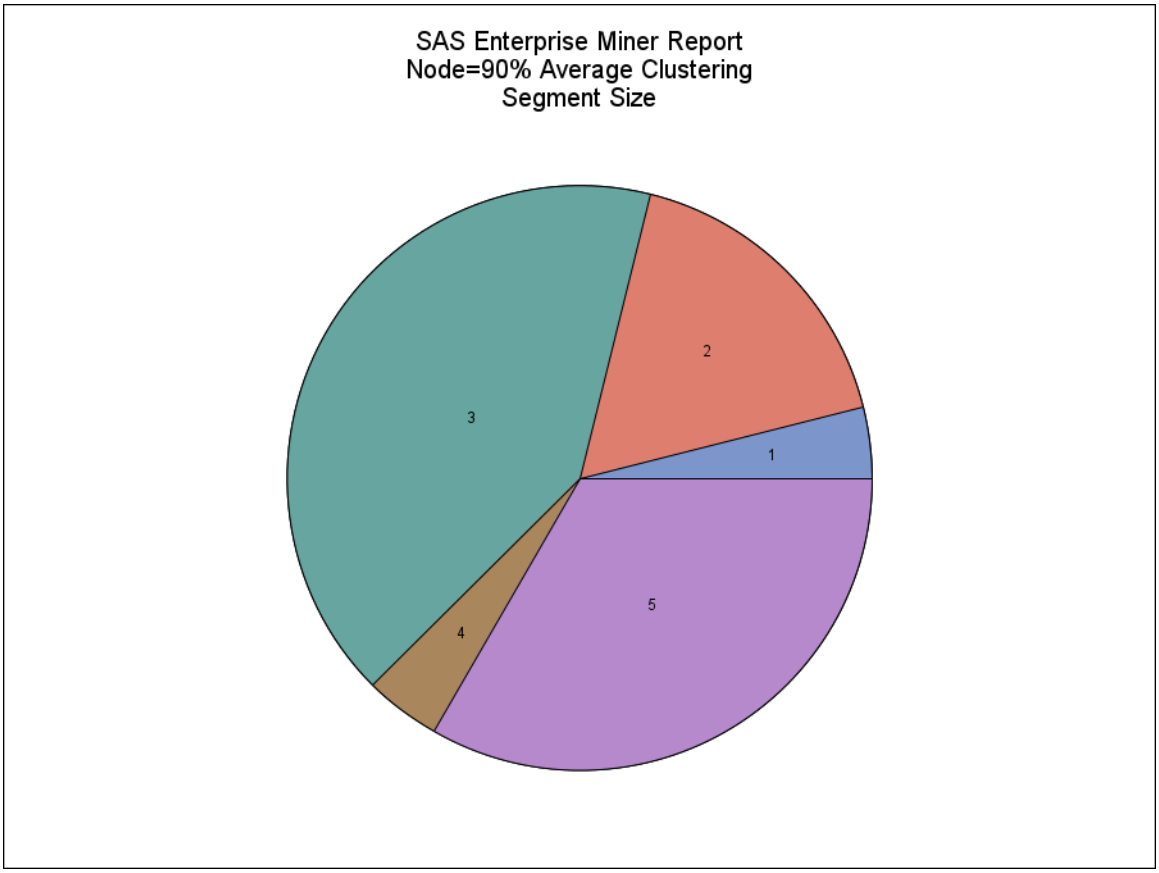
Node=90% Average Clustering Variable Summary

Role	Level	Frequency Count	Name
INPUT	BINARY	8	AnyHealthcare CholCheck DiffWalk HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth
ID	INTERVAL	1	_dataobs_

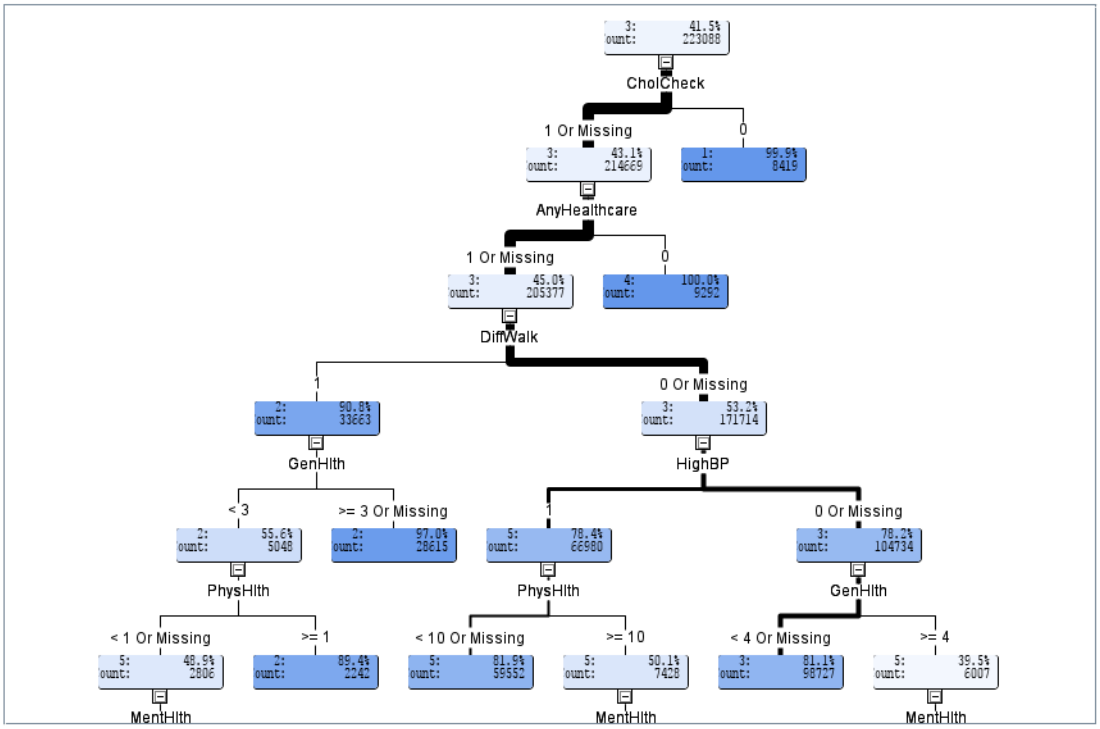
Node=90% Average Clustering Mean Statistics

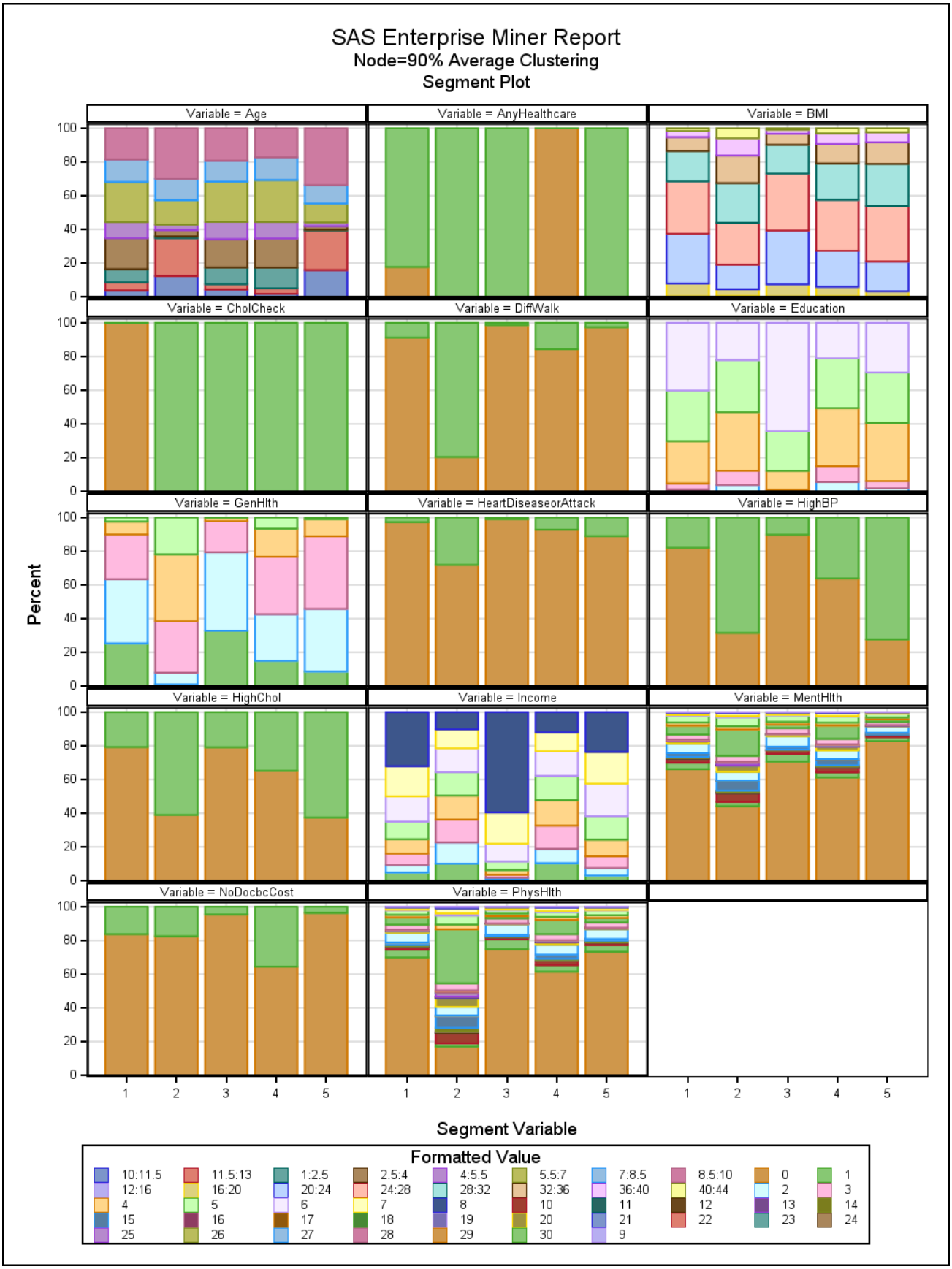
Variable	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5
Clustering Criterion	0.67	0.67	0.67	0.67	0.67
Maximum Relative Change in Cluster Seeds	0.01	0.01	0.01	0.01	0.01
Improvement in Clustering Criterion	.	.	.	.	.
Frequency of Cluster	8414.00	38384.00	92495.00	9293.00	74502.00
Root-Mean-Square Standard Deviation	0.81	0.79	0.60	0.80	0.65
Maximum Distance from Cluster Seed	8.73	7.11	7.25	8.01	7.55
Nearest Cluster	3.00	5.00	5.00	5.00	3.00
Distance to Nearest Cluster	5.41	3.02	2.33	5.11	2.33
Age	6.65	9.32	6.50	6.18	9.71
BMI	26.80	30.01	26.29	28.23	28.68
AnyHealthcare=0	0.18	0.00	0.00	1.00	0.00
AnyHealthcare=1	0.82	1.00	1.00	-0.00	1.00

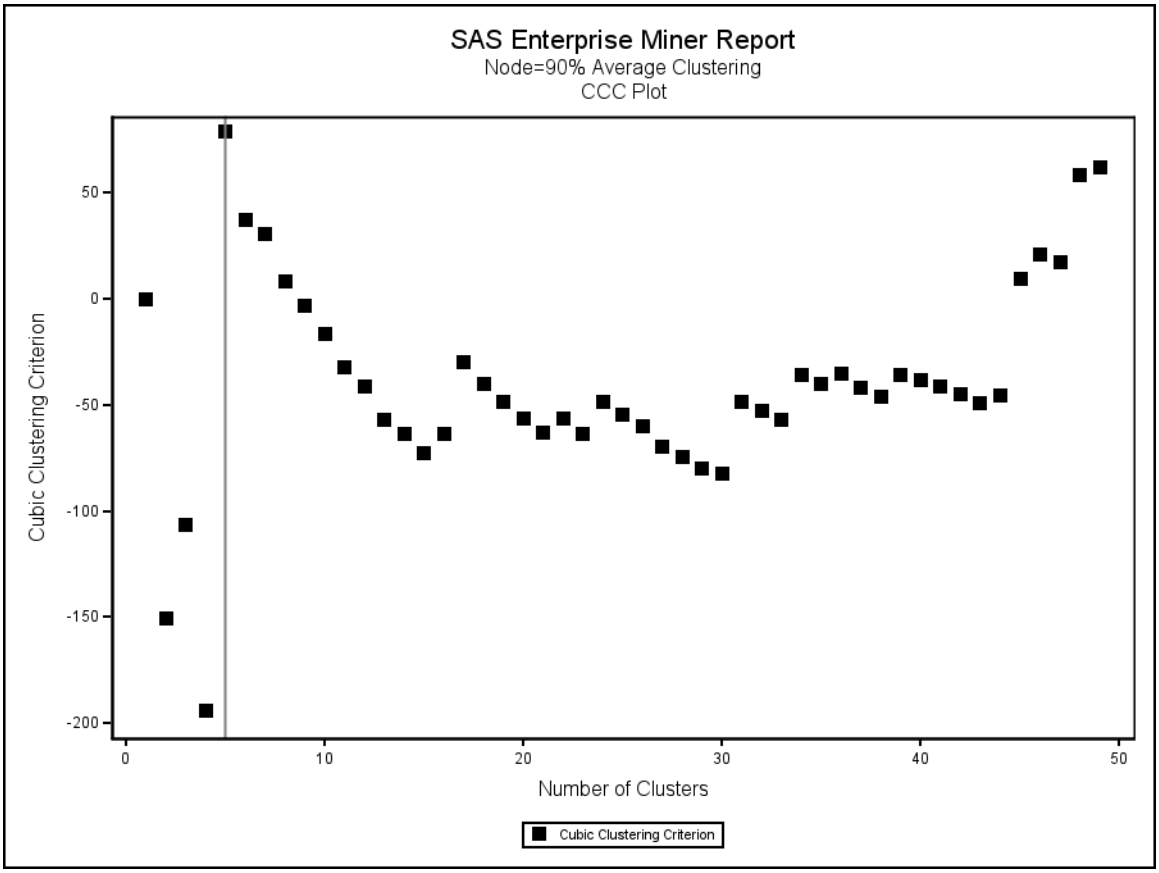
Variable	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5
CholCheck=0	1.00	0.00	0.00	0.00	-0.00
CholCheck=1	0.00	1.00	1.00	1.00	1.00
DiffWalk=0	0.91	0.20	0.99	0.84	0.97
DiffWalk=1	0.09	0.80	0.01	0.16	0.03
HeartDiseaseorAttack=0	0.97	0.72	0.99	0.93	0.89
HeartDiseaseorAttack=1	0.03	0.28	0.01	0.07	0.11
HighBP=0	0.82	0.32	0.90	0.64	0.28
HighBP=1	0.18	0.68	0.10	0.36	0.72
HighChol=0	0.79	0.39	0.79	0.65	0.37
HighChol=1	0.21	0.61	0.21	0.35	0.63
HvyAlcoholConsump=0	0.92	0.97	0.93	0.94	0.95
HvyAlcoholConsump=1	0.08	0.03	0.07	0.06	0.05
NoDocbcCost=0	0.84	0.83	0.95	0.64	0.96
NoDocbcCost=1	0.16	0.17	0.05	0.36	0.04
Transform: Education	0.49	0.37	0.63	0.36	0.42
Transform: GenHlth	0.43	0.80	0.34	0.56	0.54
Transform: Income	0.48	0.29	0.65	0.31	0.44
Transform: MentHlth	0.52	0.65	0.49	0.55	0.43
Transform: PhysHlth	0.46	0.79	0.43	0.51	0.44



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




SAS Enterprise Miner Report
Node=90% Average Clustering
Segment Plot
Variable = Age
Variable = AnyHealthcare
Variable = BMI
Variable = CholCheck
Variable = DiffWalk
Variable = Education
Variable = GenHlth
Variable = HeartDiseaseorAttack
Variable = HighBP
Variable = HighChol
Variable = Income
Variable = MentHlth
Variable = NoDocbcCost
Variable = PhysHlth
Percent
0
20
40
60
80
100
1
2
3
4
5
Segment Variable
Formatted Value
10:11.5
12:16
4
15
25
11.5:13
16:20
5
16
26
1:2.5
20:24
6
17
27
2.5:4
24:28
7
18
28
4:5.5
28:32
8
19
29
5.5:7
32:36
10
20
30
7:8.5
36:40
11
21
9
8.5:10
40:44
12
22
0
2
13
23
1
3
14
24


SAS Enterprise Miner Report
Node=90% Average Clustering
CCC Plot
Cubic Clustering Criterion
50
0
-50
-100
-150
-200
0
10
20
30
40
50
Number of Clusters
Cubic Clustering Criterion

SAS Enterprise Miner Report

Node=90% Centroid Clustering Summary

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

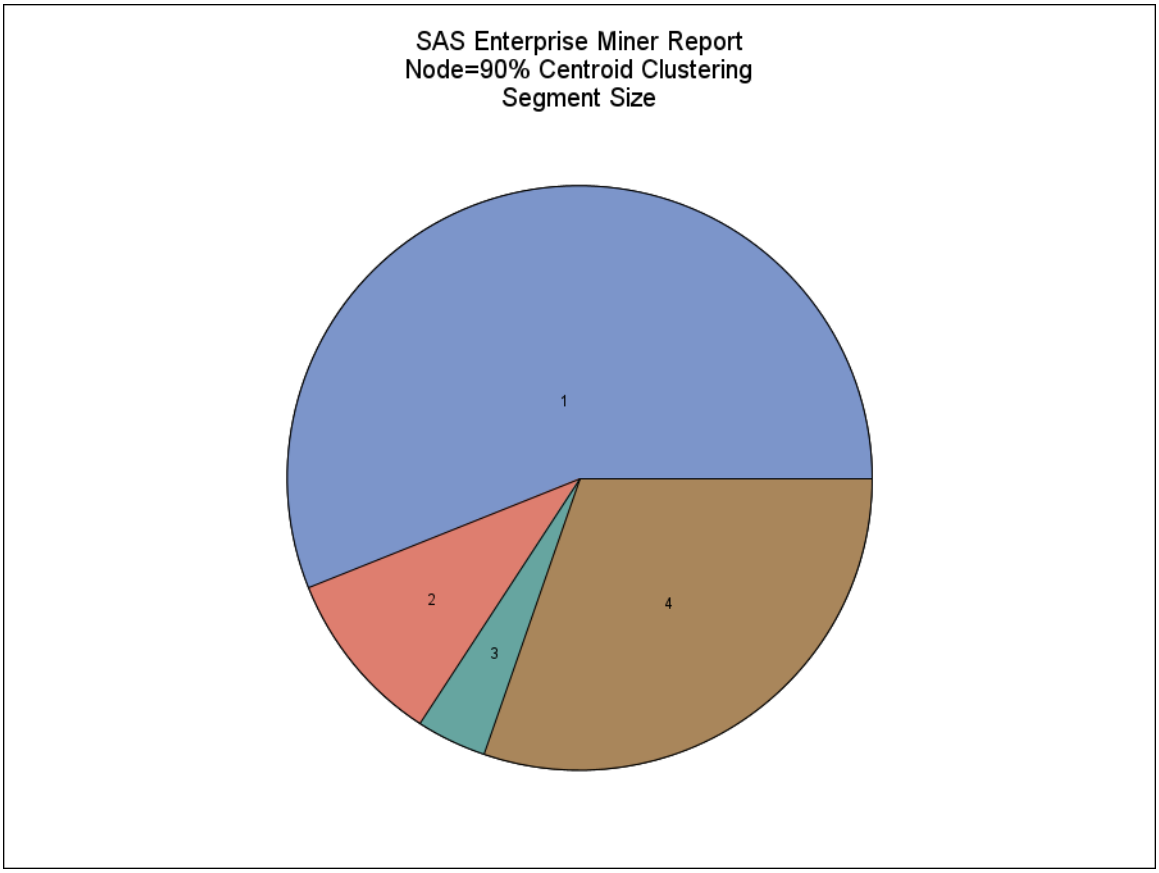
Node=90% Centroid Clustering Variable Summary

Role	Level	Frequency Count	Name
INPUT	BINARY	8	AnyHealthcare CholCheck DiffWalk HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth
ID	INTERVAL	1	_dataobs_

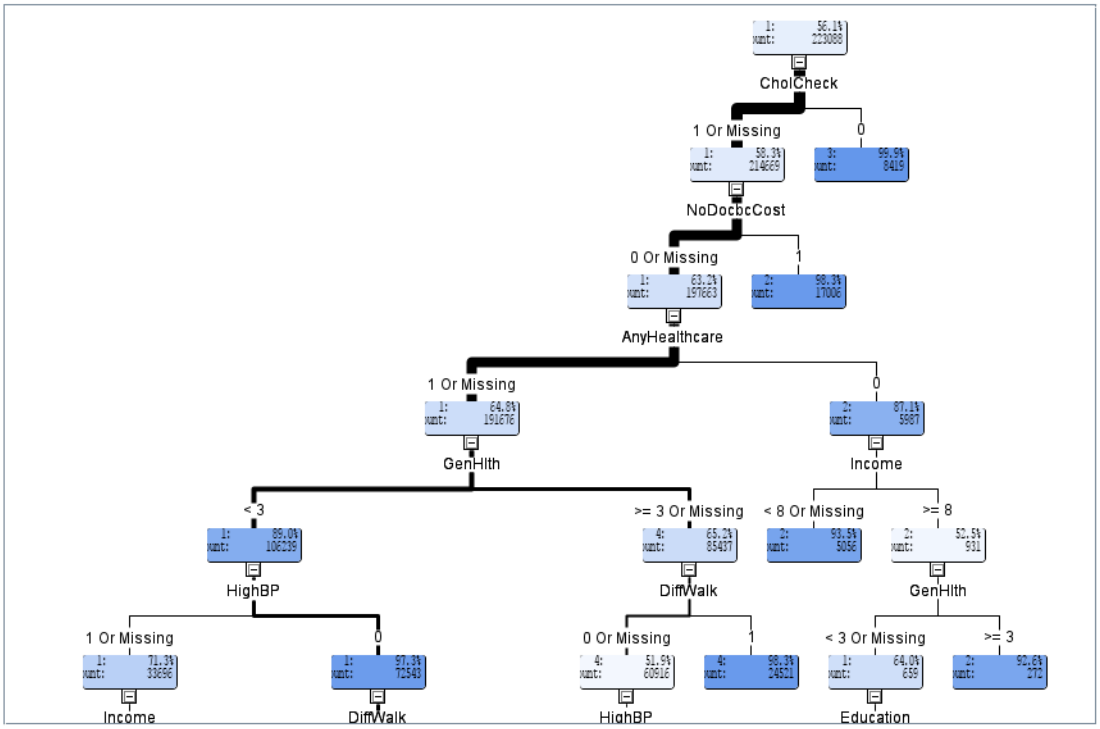
Node=90% Centroid Clustering Mean Statistics

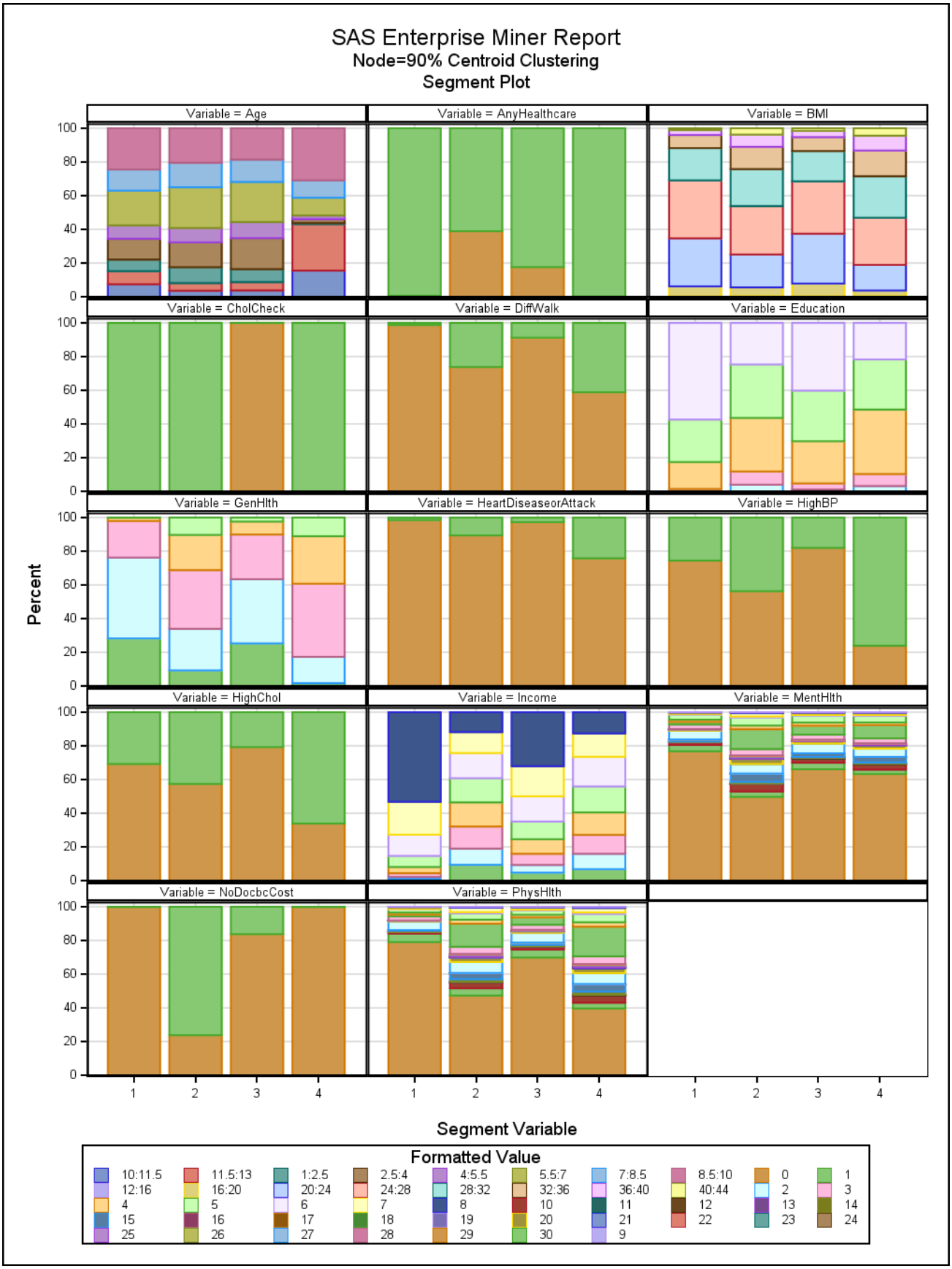
Variable	Segment 1	Segment 2	Segment 3	Segment 4
Clustering Criterion	0.70	0.70	0.70	0.70
Maximum Relative Change in Cluster Seeds	0.00	0.00	0.00	0.00
Improvement in Clustering Criterion	.	.	.	.
Frequency of Cluster	125140.00	21930.00	8414.00	67604.00
Root-Mean-Square Standard Deviation	0.60	0.94	0.81	0.74
Maximum Distance from Cluster Seed	7.00	7.90	8.73	6.85
Nearest Cluster	4.00	4.00	1.00	1.00
Distance to Nearest Cluster	2.93	3.65	5.41	2.93
Age	7.40	6.72	6.65	9.85
BMI	26.80	28.71	26.80	29.59
AnyHealthcare=0	0.01	0.39	0.18	0.00
AnyHealthcare=1	0.99	0.61	0.82	1.00

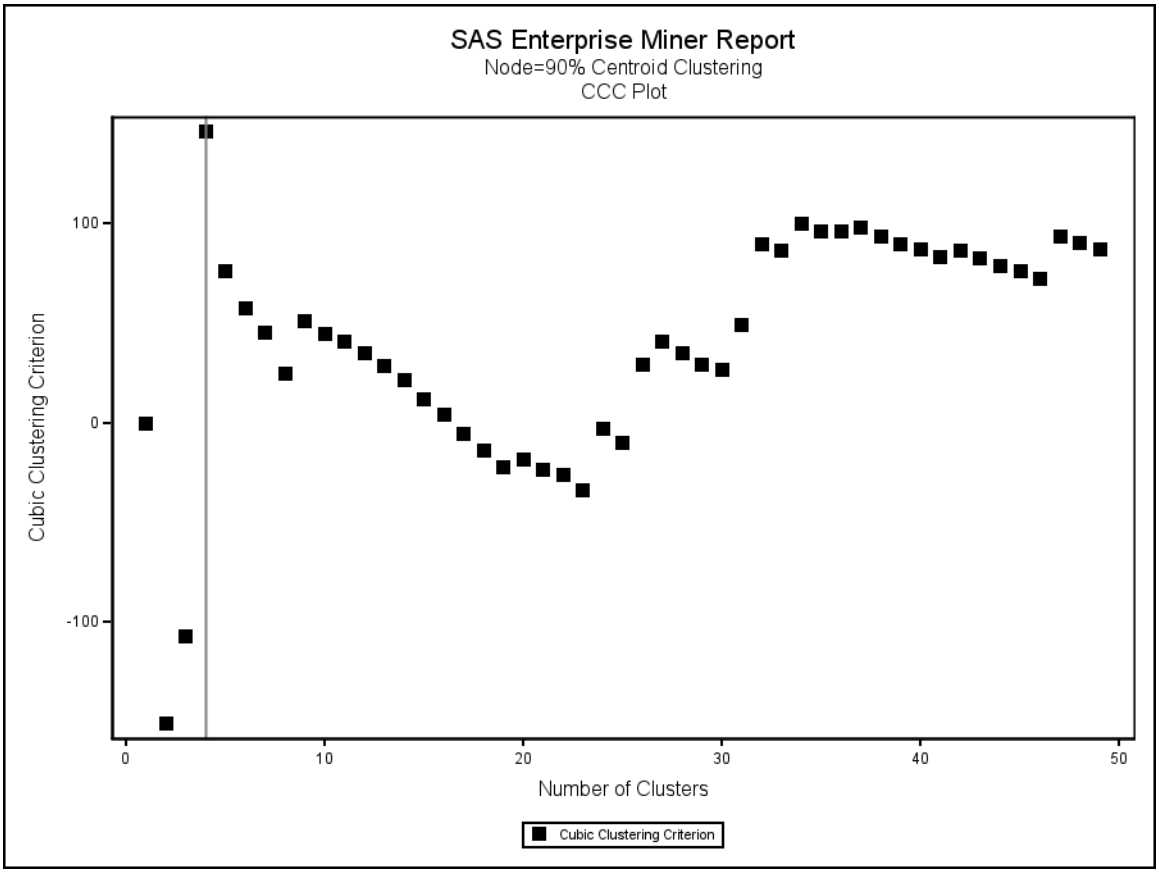
Variable	Segment 1	Segment 2	Segment 3	Segment 4
CholCheck=0	0.00	0.00	1.00	-0.00
CholCheck=1	1.00	1.00	0.00	1.00
DiffWalk=0	0.99	0.74	0.91	0.59
DiffWalk=1	0.01	0.26	0.09	0.41
HeartDiseaseorAttack=0	0.99	0.89	0.97	0.76
HeartDiseaseorAttack=1	0.01	0.11	0.03	0.24
HighBP=0	0.74	0.56	0.82	0.24
HighBP=1	0.26	0.44	0.18	0.76
HighChol=0	0.69	0.57	0.79	0.34
HighChol=1	0.31	0.43	0.21	0.66
HvyAlcoholConsump=0	0.93	0.94	0.92	0.97
HvyAlcoholConsump=1	0.07	0.06	0.08	0.03
NoDocbcCost=0	1.00	0.24	0.84	1.00
NoDocbcCost=1	0.00	0.76	0.16	0.00
Transform: Education	0.59	0.39	0.49	0.37
Transform: GenHlth	0.37	0.62	0.43	0.72
Transform: Income	0.62	0.31	0.48	0.34
Transform: MentHlth	0.46	0.61	0.52	0.54
Transform: PhysHlth	0.41	0.59	0.46	0.64



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




SAS Enterprise Miner Report
Node=90% Centroid Clustering
Segment Plot
Variable = Age
Variable = AnyHealthcare
Variable = BMI
Variable = CholCheck
Variable = DiffWalk
Variable = Education
Variable = GenHlth
Variable = HeartDiseaseorAttack
Variable = HighBP
Variable = HighChol
Variable = Income
Variable = MentHlth
Variable = NoDocbcCost
Variable = PhysHlth
Percent
0
20
40
60
80
100
1
2
3
4
Segment Variable
Formatted Value
10:11.5
12:16
4
15
25
11.5:13
16:20
5
16
26
1:2.5
20:24
6
17
27
2.5:4
24:28
7
18
28
4:5.5
28:32
8
19
29
5.5:7
32:36
10
20
30
7:8.5
36:40
11
21
9
8.5:10
40:44
12
22
0
2
13
23
1
3
14
24


SAS Enterprise Miner Report
Node=90% Centroid Clustering
CCC Plot
Cubic Clustering Criterion
100
0
-100
0
10
20
30
40
50
Number of Clusters
Cubic Clustering Criterion

SAS Enterprise Miner Report

Node=90% Ward Clustering Summary

Node id = Clus4
Node label = 90% Ward Clustering
Meta path = Ids => Filter => Smpl => 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				

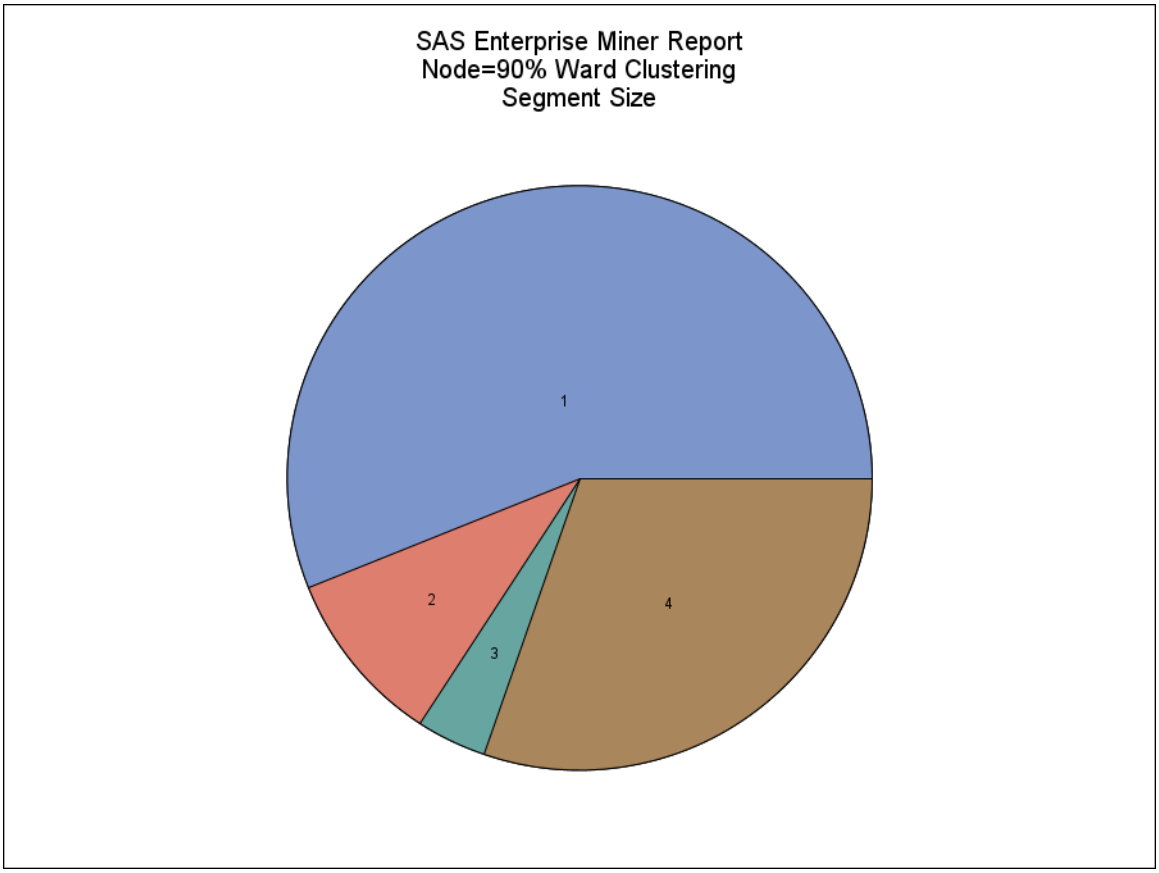
Node=90% Ward Clustering Variable Summary

Role	Level	Frequency Count	Name
INPUT	BINARY	8	AnyHealthcare CholCheck DiffWalk HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth
ID	INTERVAL	1	_dataobs_

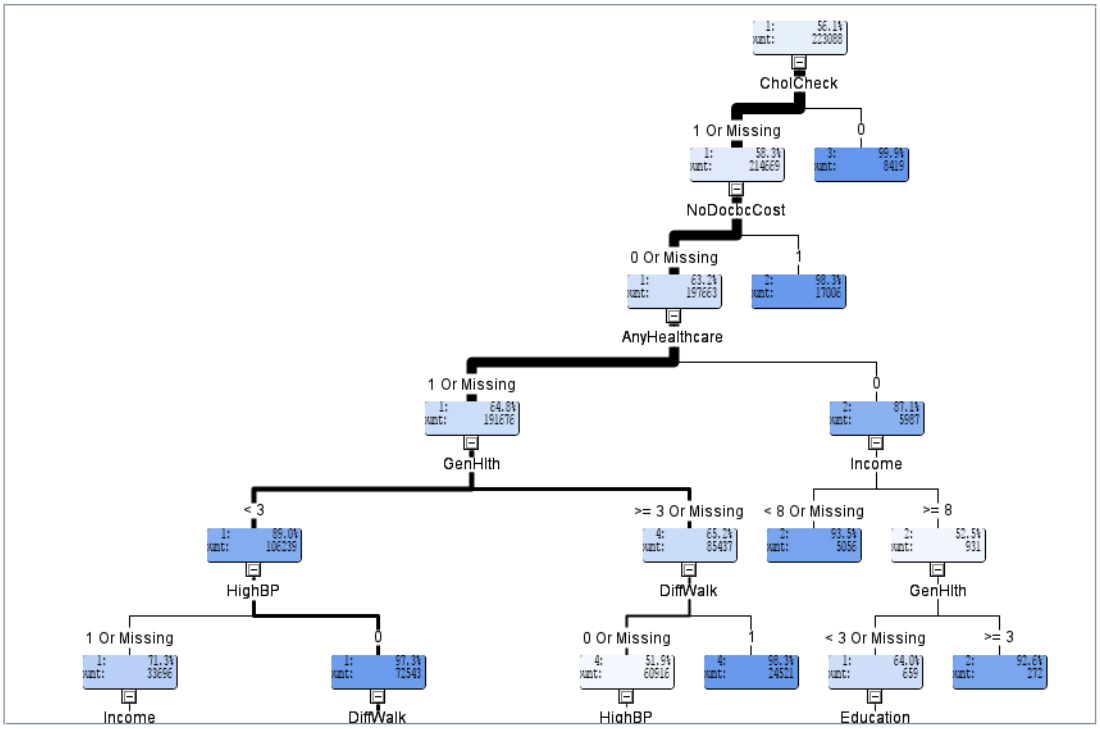
Node=90% Ward Clustering Mean Statistics

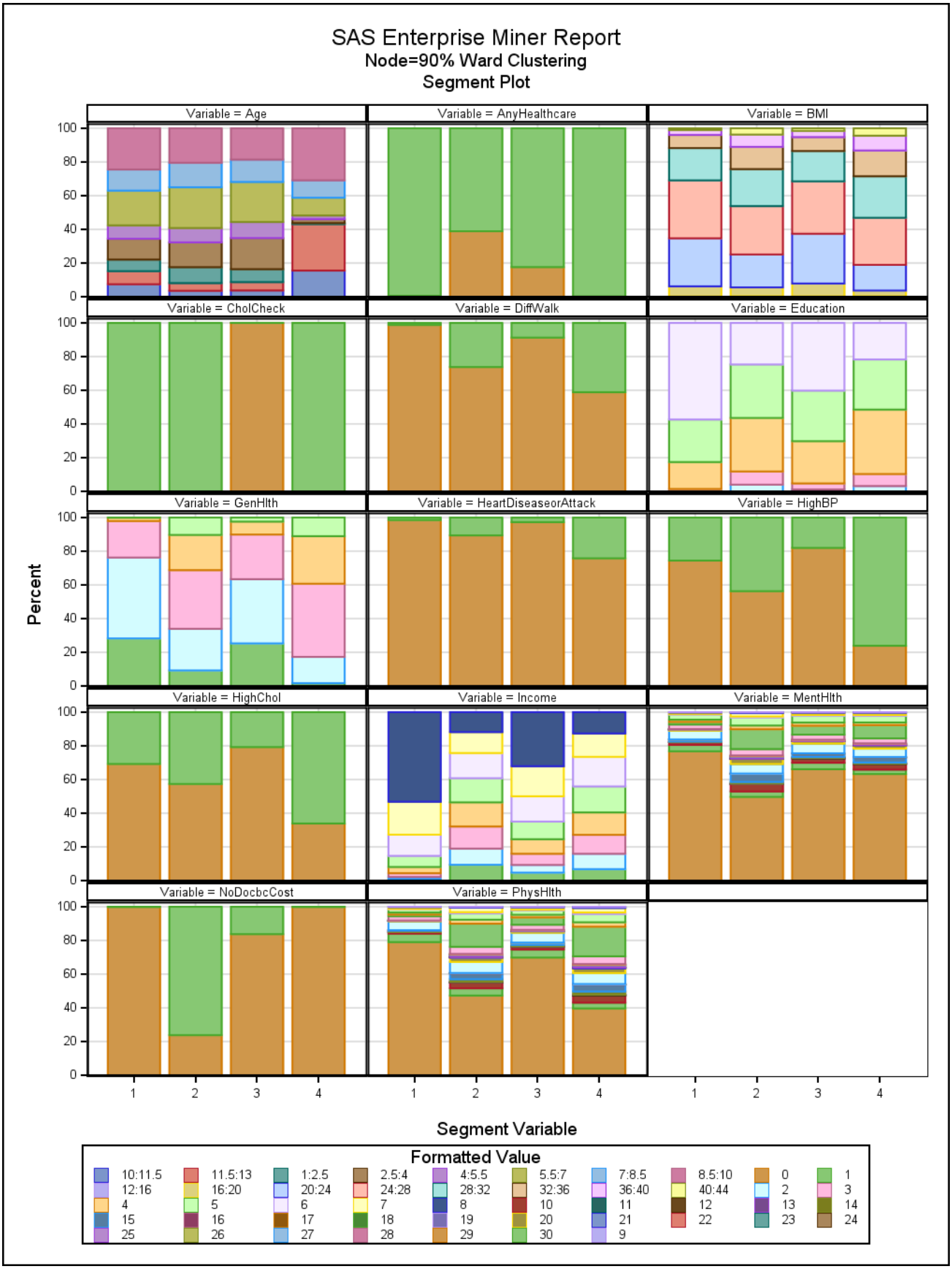
Variable	Segment 1	Segment 2	Segment 3	Segment 4
Clustering Criterion	0.70	0.70	0.70	0.70
Maximum Relative Change in Cluster Seeds	0.00	0.00	0.00	0.00
Improvement in Clustering Criterion	.	.	.	.
Frequency of Cluster	125140.00	21930.00	8414.00	67604.00
Root-Mean-Square Standard Deviation	0.60	0.94	0.81	0.74
Maximum Distance from Cluster Seed	7.00	7.90	8.73	6.85
Nearest Cluster	4.00	4.00	1.00	1.00
Distance to Nearest Cluster	2.93	3.65	5.41	2.93
Age	7.40	6.72	6.65	9.85
BMI	26.80	28.71	26.80	29.59
AnyHealthcare=0	0.01	0.39	0.18	0.00
AnyHealthcare=1	0.99	0.61	0.82	1.00

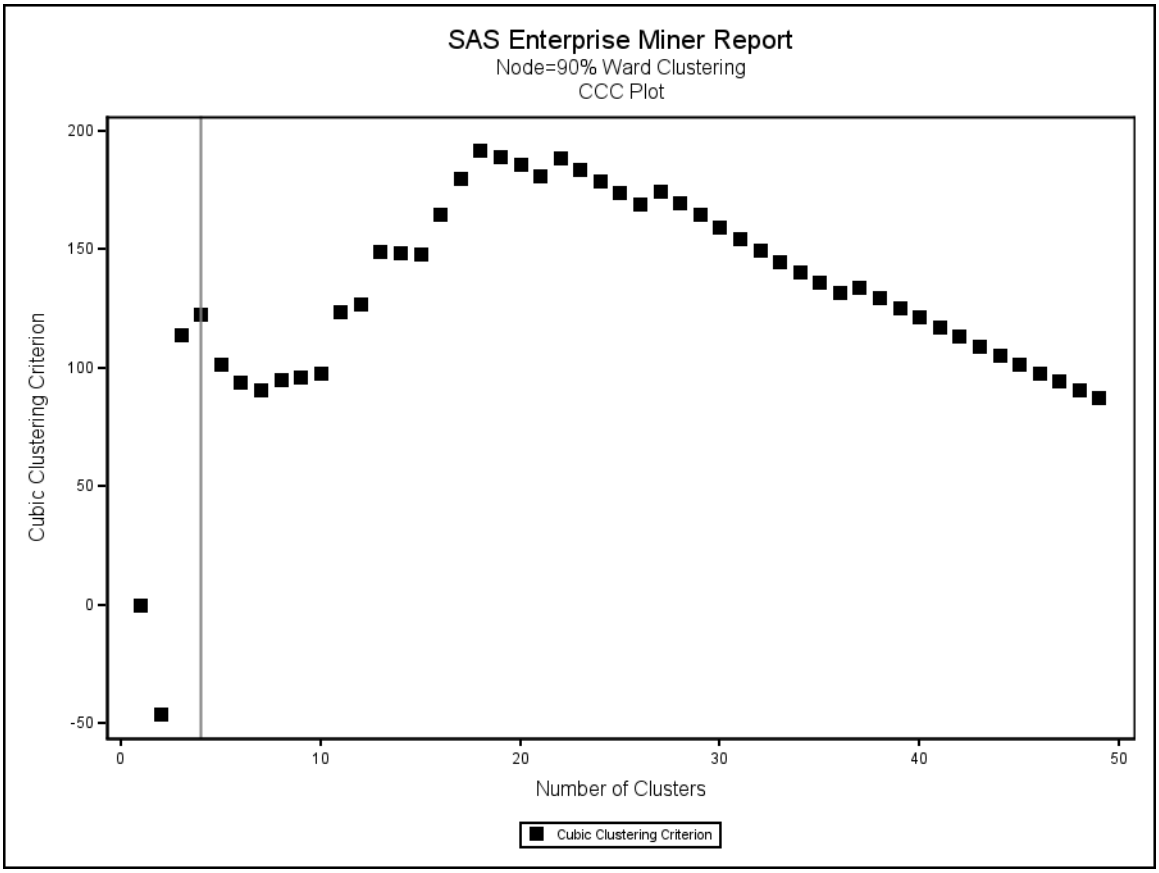
Variable	Segment 1	Segment 2	Segment 3	Segment 4
CholCheck=0	0.00	0.00	1.00	-0.00
CholCheck=1	1.00	1.00	0.00	1.00
DiffWalk=0	0.99	0.74	0.91	0.59
DiffWalk=1	0.01	0.26	0.09	0.41
HeartDiseaseorAttack=0	0.99	0.89	0.97	0.76
HeartDiseaseorAttack=1	0.01	0.11	0.03	0.24
HighBP=0	0.74	0.56	0.82	0.24
HighBP=1	0.26	0.44	0.18	0.76
HighChol=0	0.69	0.57	0.79	0.34
HighChol=1	0.31	0.43	0.21	0.66
HvyAlcoholConsump=0	0.93	0.94	0.92	0.97
HvyAlcoholConsump=1	0.07	0.06	0.08	0.03
NoDocbcCost=0	1.00	0.24	0.84	1.00
NoDocbcCost=1	0.00	0.76	0.16	0.00
Transform: Education	0.59	0.39	0.49	0.37
Transform: GenHlth	0.37	0.62	0.43	0.72
Transform: Income	0.62	0.31	0.48	0.34
Transform: MentHlth	0.46	0.61	0.52	0.54
Transform: PhysHlth	0.41	0.59	0.46	0.64



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




SAS Enterprise Miner Report
Node=90% Ward Clustering
Segment Plot
Variable = Age
Variable = AnyHealthcare
Variable = BMI
Variable = CholCheck
Variable = DiffWalk
Variable = Education
Variable = GenHlth
Variable = HeartDiseaseorAttack
Variable = HighBP
Variable = HighChol
Variable = Income
Variable = MentHlth
Variable = NoDocbcCost
Variable = PhysHlth
Percent
0
20
40
60
80
100
1
2
3
4
Segment Variable
Formatted Value
10:11.5
11.5:13
1:2.5
2.5:4
4:5.5
5.5:7
7:8.5
8.5:10
0
1
12:16
16:20
20:24
24:28
28:32
32:36
36:40
40:44
2
3
4
5
6
7
8
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
9


SAS Enterprise Miner Report
Node=90% Ward Clustering
CCC Plot
Cubic Clustering Criterion
200
150
100
50
0
-50
0
10
20
30
40
50
Number of Clusters
Cubic Clustering Criterion

SAS Enterprise Miner Report

Node=Average Clustering Summary

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

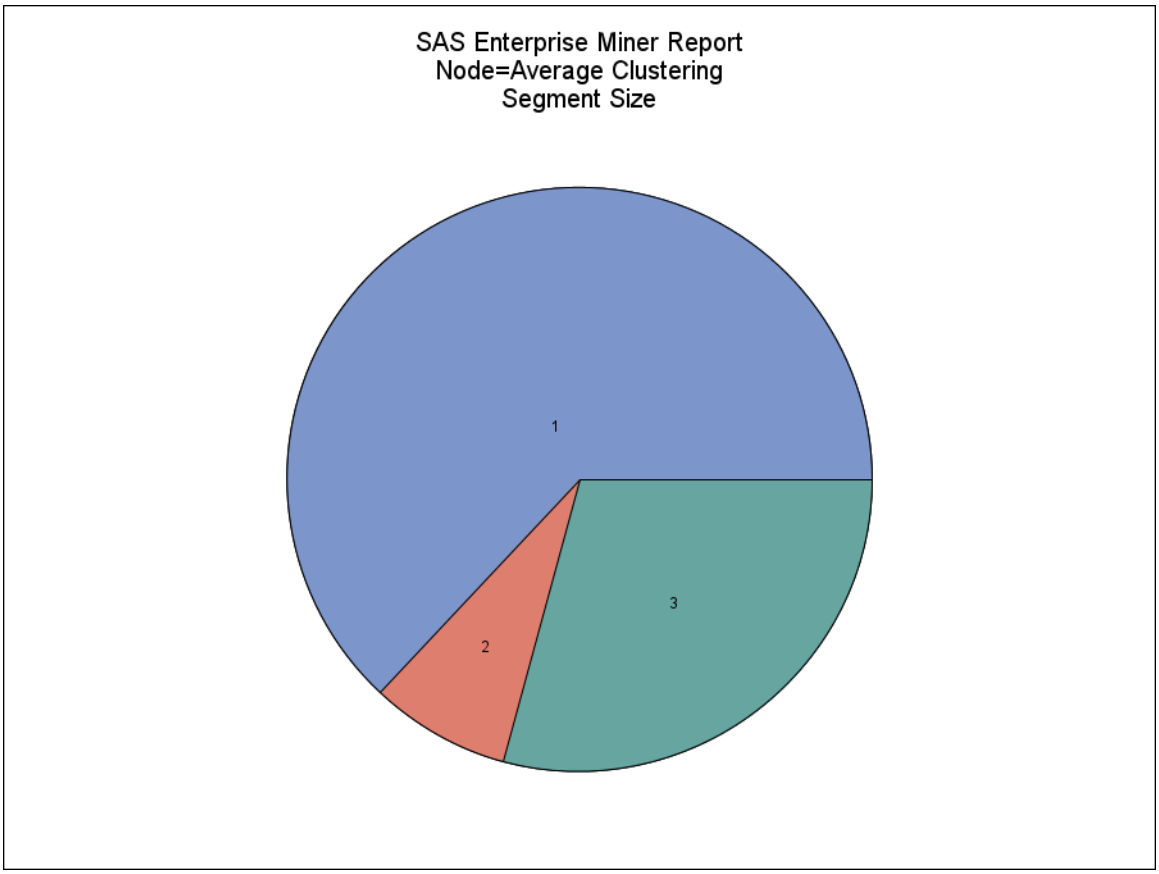
Node=Average Clustering Variable Summary

Role	Level	Frequency Count	Name
INPUT	BINARY	8	AnyHealthcare CholCheck DiffWalk HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth

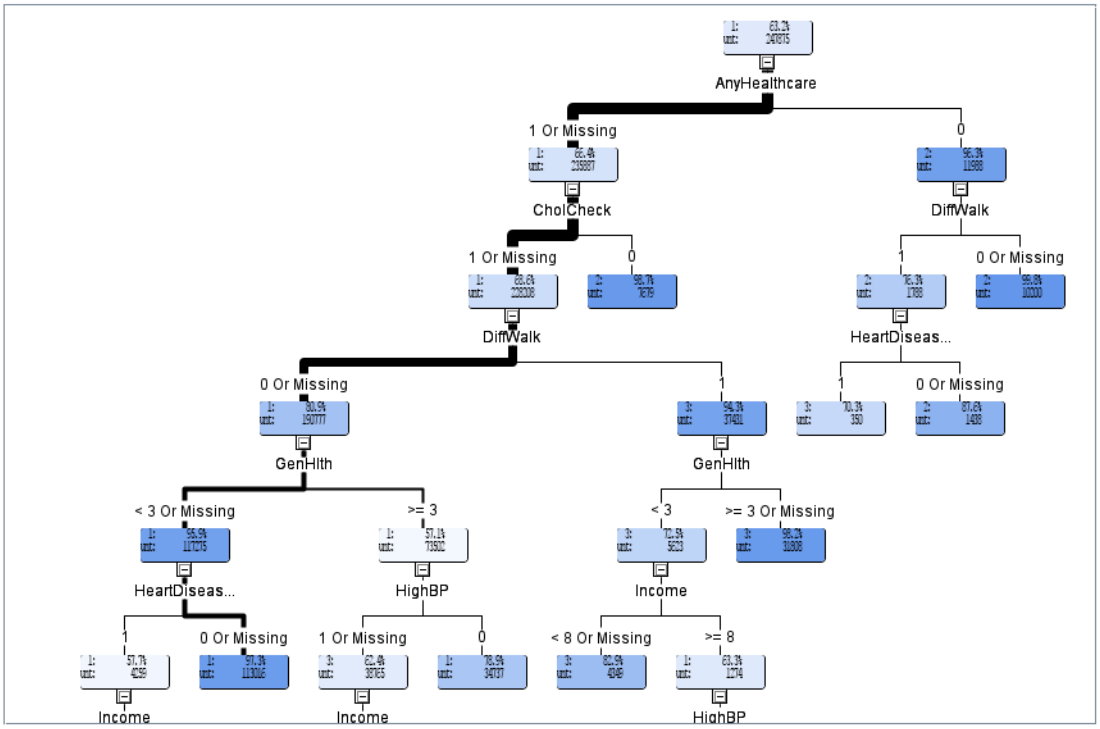
Node=Average Clustering Mean Statistics

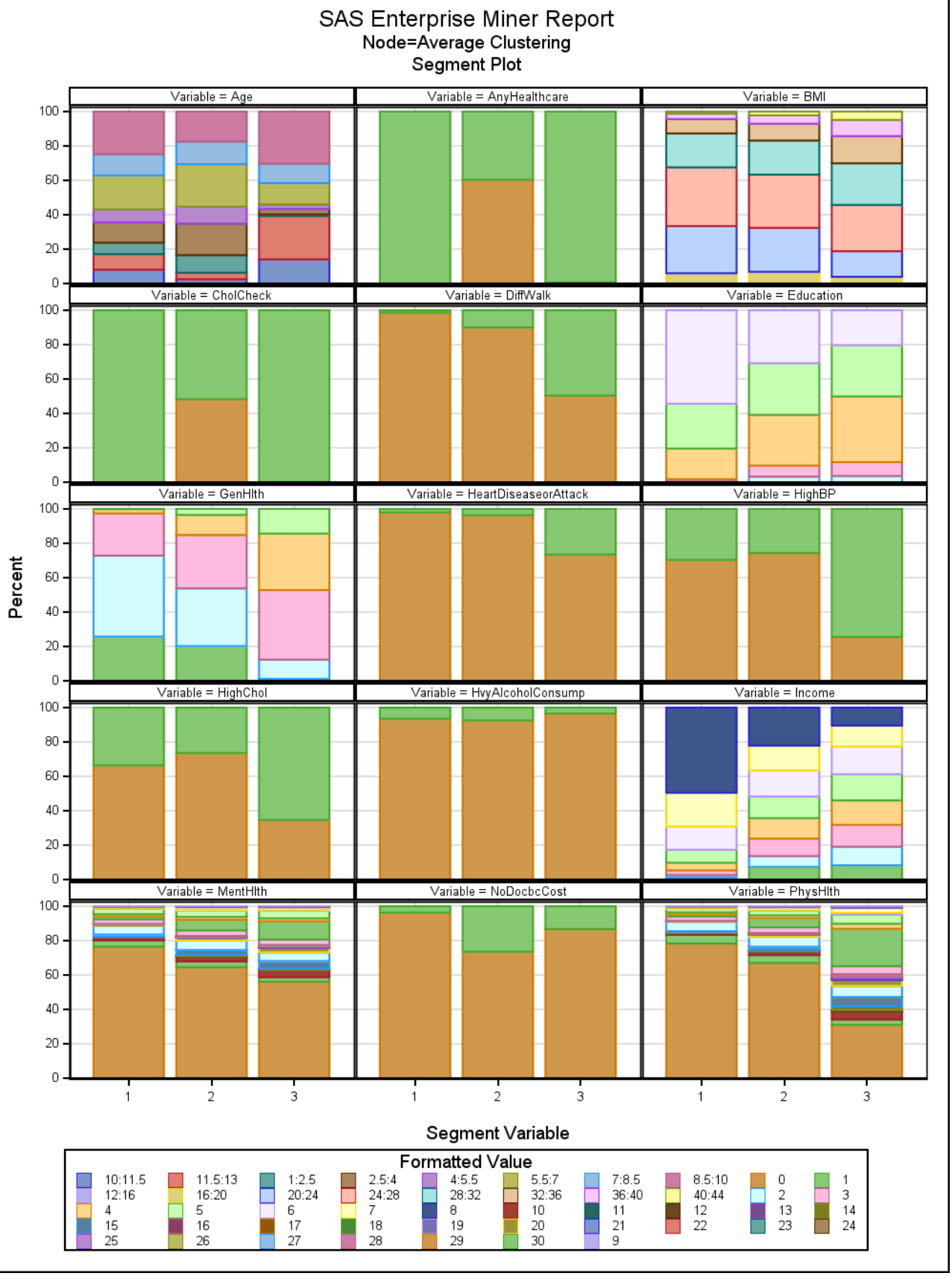
Variable	Segment 1	Segment 2	Segment 3
Clustering Criterion	0.72	0.72	0.72
Maximum Relative Change in Cluster Seeds	0.00	0.00	0.00
Improvement in Clustering Criterion	.	.	.
Frequency of Cluster	156554.00	19127.00	72194.00
Root-Mean-Square Standard Deviation	0.63	1.05	0.80
Maximum Distance from Cluster Seed	7.02	8.22	8.23
Nearest Cluster	3.00	1.00	1.00
Distance to Nearest Cluster	3.09	4.05	3.09
Age	7.55	6.34	9.59
BMI	26.98	27.47	29.75
AnyHealthcare=0	0.00	0.60	0.01
AnyHealthcare=1	1.00	0.40	0.99
CholCheck=0	0.00	0.48	0.00

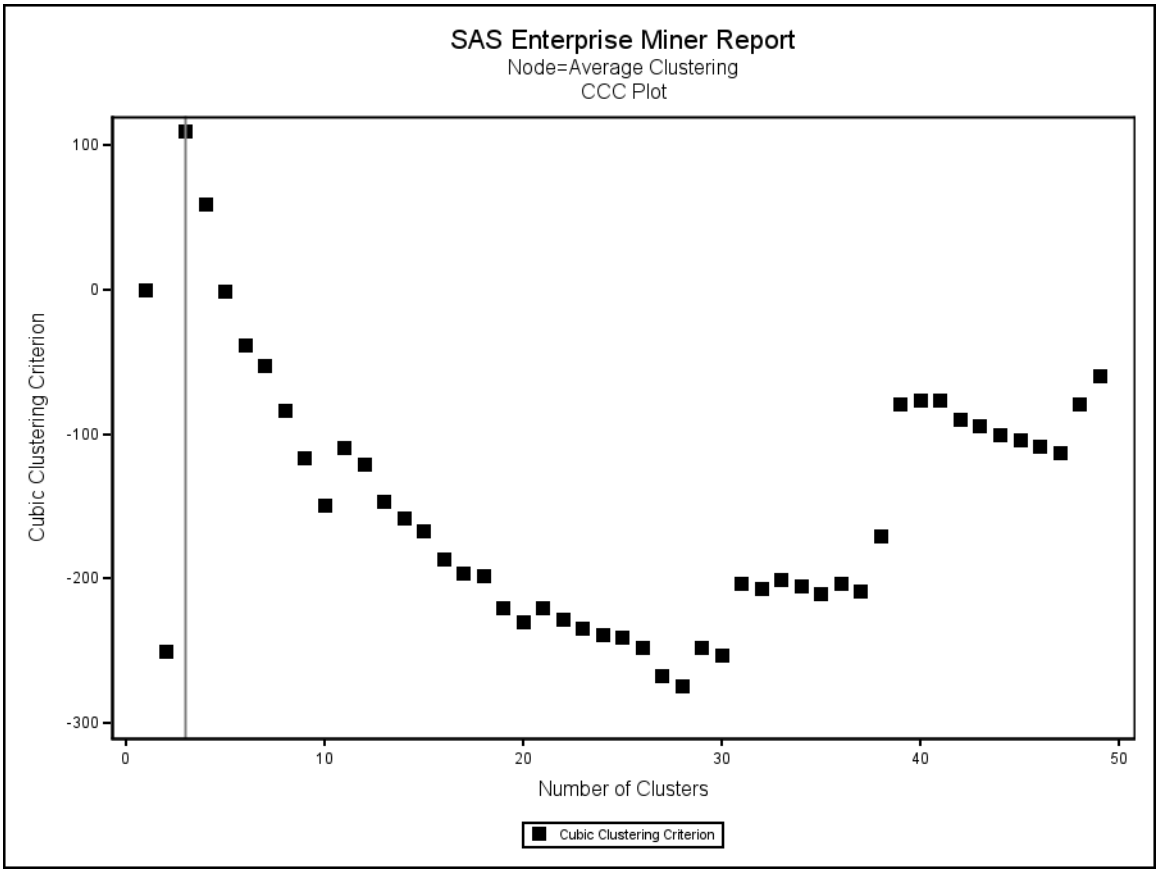
Variable	Segment 1	Segment 2	Segment 3
CholCheck=1	1.00	0.52	1.00
DiffWalk=0	0.99	0.90	0.50
DiffWalk=1	0.01	0.10	0.50
HeartDiseaseorAttack=0	0.98	0.96	0.73
HeartDiseaseorAttack=1	0.02	0.04	0.27
HighBP=0	0.70	0.74	0.26
HighBP=1	0.30	0.26	0.74
HighChol=0	0.66	0.74	0.35
HighChol=1	0.34	0.26	0.65
HvyAlcoholConsump=0	0.94	0.93	0.97
HvyAlcoholConsump=1	0.06	0.07	0.03
NoDocbcCost=0	0.96	0.74	0.87
NoDocbcCost=1	0.04	0.26	0.13
Transform: Education	0.57	0.43	0.36
Transform: GenHlth	0.38	0.49	0.75
Transform: Income	0.60	0.39	0.31
Transform: MentHlth	0.46	0.53	0.58
Transform: PhysHlth	0.41	0.48	0.69



SAS Enterprise Miner Report
Node=Average Clustering
Tree Diagram




SAS Enterprise Miner Report
Node=Average Clustering
Segment Plot
Variable = Age
Variable = AnyHealthcare
Variable = BMI
Variable = CholCheck
Variable = DiffWalk
Variable = Education
Variable = GenHlth
Variable = HeartDiseaseorAttack
Variable = HighBP
Variable = HighChol
Variable = HvyAlcoholConsump
Variable = Income
Variable = MentHlth
Variable = NoDocbcCost
Variable = PhysHlth
Percent
0
20
40
60
80
100
1
2
3
Segment Variable
Formatted Value
10:11.5
11.5:13
1:2.5
2.5:4
4:5.5
5.5:7
7:8.5
8.5:10
0
1
12:16
16:20
20:24
24:28
28:32
32:36
36:40
40:44
2
3
4
5
6
7
8
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
9


SAS Enterprise Miner Report
Node=Average Clustering
CCC Plot
Cubic Clustering Criterion
100
0
-100
-200
-300
0
10
20
30
40
50
Number of Clusters
Cubic Clustering Criterion

SAS Enterprise Miner Report

Node=Centroid Clustering Summary

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

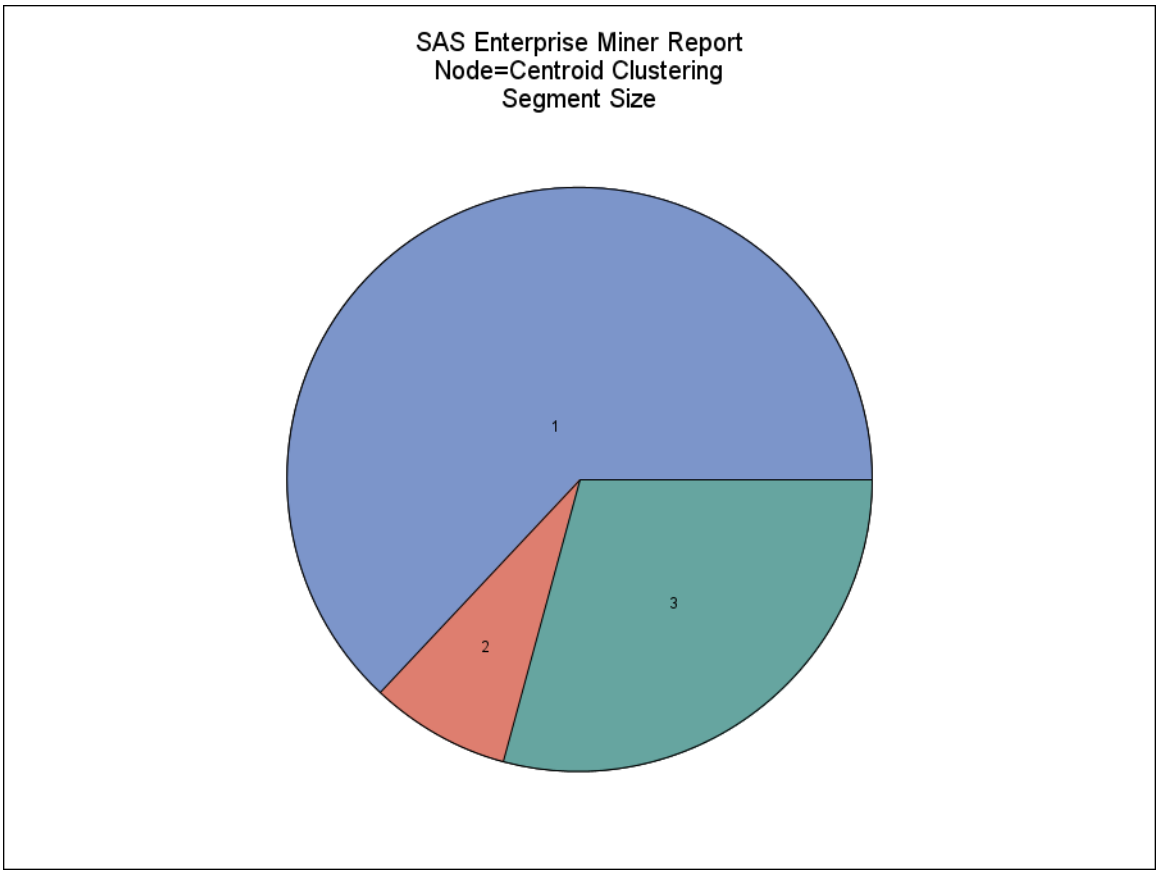
Node=Centroid Clustering Variable Summary

Role	Level	Frequency Count	Name
INPUT	BINARY	8	AnyHealthcare CholCheck DiffWalk HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth

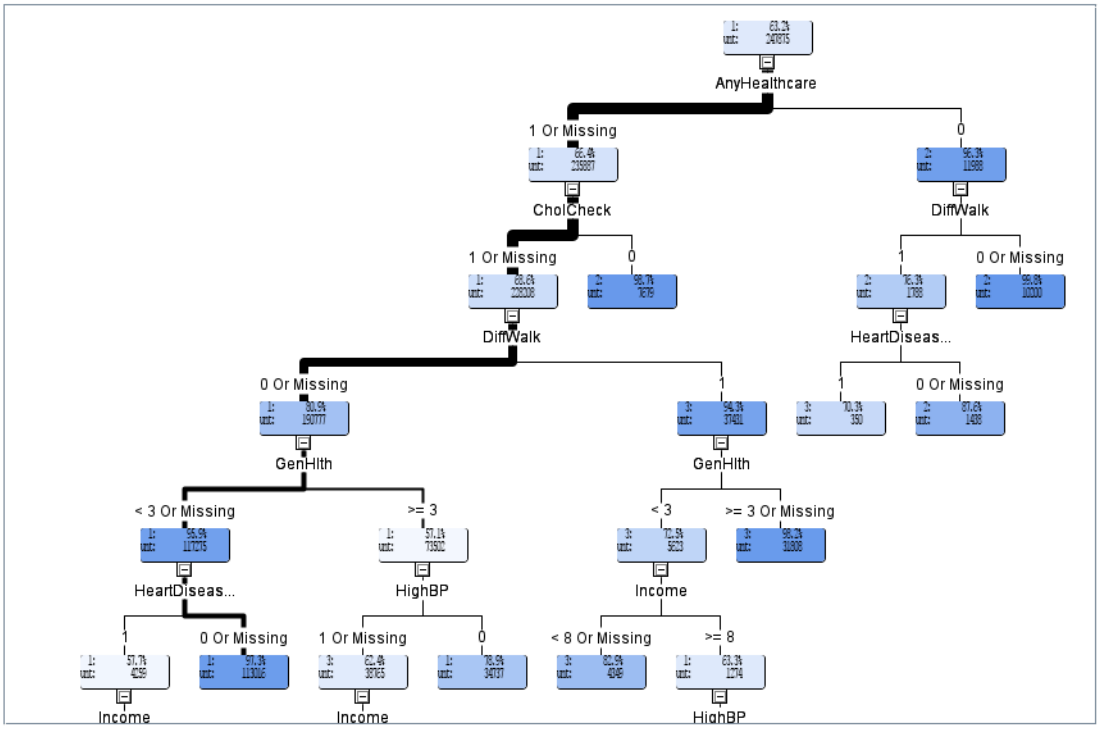
Node=Centroid Clustering Mean Statistics

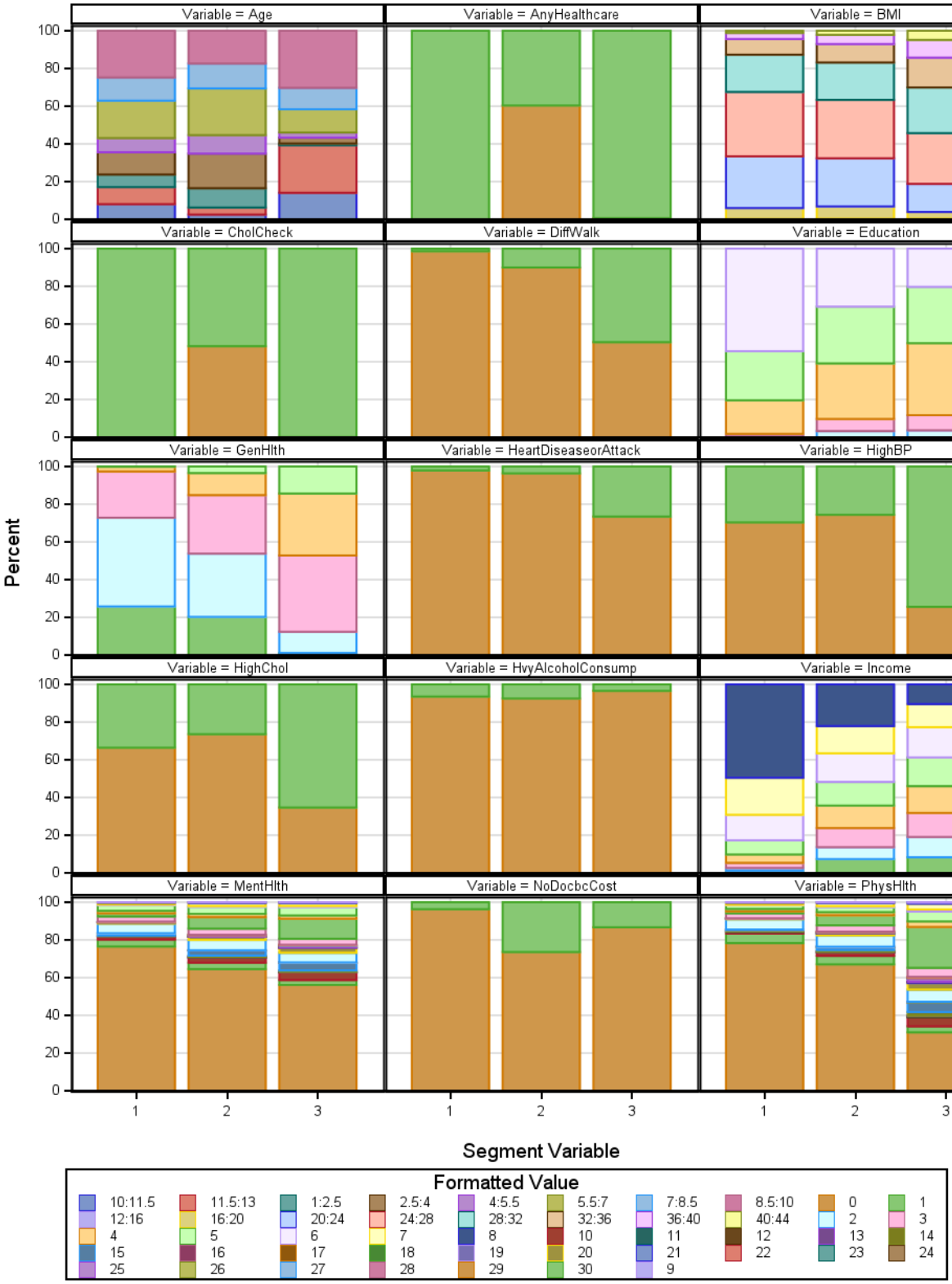
Variable	Segment 1	Segment 2	Segment 3
Clustering Criterion	0.72	0.72	0.72
Maximum Relative Change in Cluster Seeds	0.00	0.00	0.00
Improvement in Clustering Criterion	.	.	.
Frequency of Cluster	156554.00	19127.00	72194.00
Root-Mean-Square Standard Deviation	0.63	1.05	0.80
Maximum Distance from Cluster Seed	7.02	8.22	8.23
Nearest Cluster	3.00	1.00	1.00
Distance to Nearest Cluster	3.09	4.05	3.09
Age	7.55	6.34	9.59
BMI	26.98	27.47	29.75
AnyHealthcare=0	0.00	0.60	0.01
AnyHealthcare=1	1.00	0.40	0.99
CholCheck=0	0.00	0.48	0.00

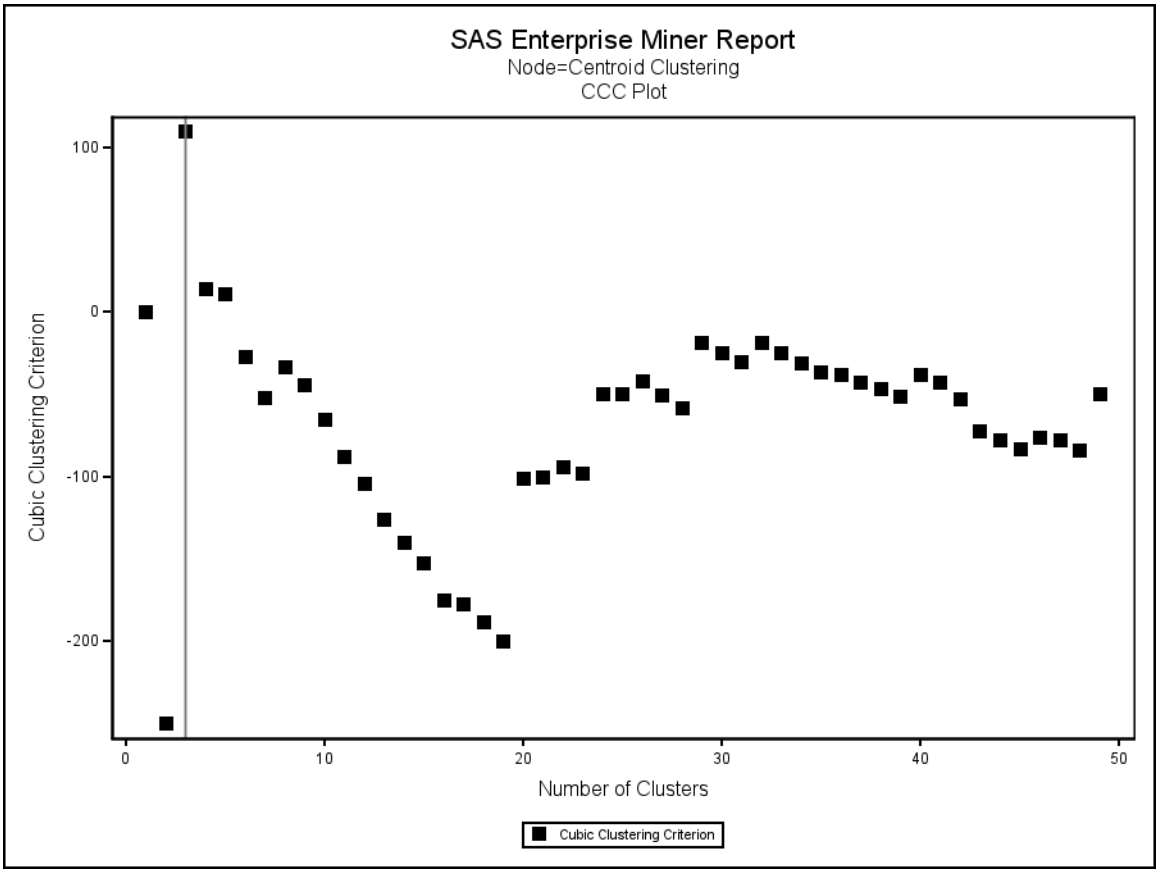
Variable	Segment 1	Segment 2	Segment 3
CholCheck=1	1.00	0.52	1.00
DiffWalk=0	0.99	0.90	0.50
DiffWalk=1	0.01	0.10	0.50
HeartDiseaseorAttack=0	0.98	0.96	0.73
HeartDiseaseorAttack=1	0.02	0.04	0.27
HighBP=0	0.70	0.74	0.26
HighBP=1	0.30	0.26	0.74
HighChol=0	0.66	0.74	0.35
HighChol=1	0.34	0.26	0.65
HvyAlcoholConsump=0	0.94	0.93	0.97
HvyAlcoholConsump=1	0.06	0.07	0.03
NoDocbcCost=0	0.96	0.74	0.87
NoDocbcCost=1	0.04	0.26	0.13
Transform: Education	0.57	0.43	0.36
Transform: GenHlth	0.38	0.49	0.75
Transform: Income	0.60	0.39	0.31
Transform: MentHlth	0.46	0.53	0.58
Transform: PhysHlth	0.41	0.48	0.69



SAS Enterprise Miner Report
Node=Centroid Clustering
Tree Diagram




SAS Enterprise Miner Report
Node=Centroid Clustering
Segment Plot
Variable = Age
Variable = AnyHealthcare
Variable = BMI
Variable = CholCheck
Variable = DiffWalk
Variable = Education
Variable = GenHlth
Variable = HeartDiseaseorAttack
Variable = HighBP
Variable = HighChol
Variable = HvyAlcoholConsump
Variable = Income
Variable = MentHlth
Variable = NoDocbcCost
Variable = PhysHlth
Percent
0
20
40
60
80
100
1
2
3
Segment Variable
Formatted Value
10:11.5
11.5:13
1:2.5
2.5:4
4:5.5
5.5:7
7:8.5
8.5:10
0
1
12:16
16:20
20:24
24:28
28:32
32:36
36:40
40:44
2
3
4
5
6
7
8
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
9


SAS Enterprise Miner Report
Node=Centroid Clustering
CCC Plot
Cubic Clustering Criterion
100
0
-100
-200
0
10
20
30
40
50
Number of Clusters
Cubic Clustering Criterion

SAS Enterprise Miner Report

Node=Ward Clustering Summary

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

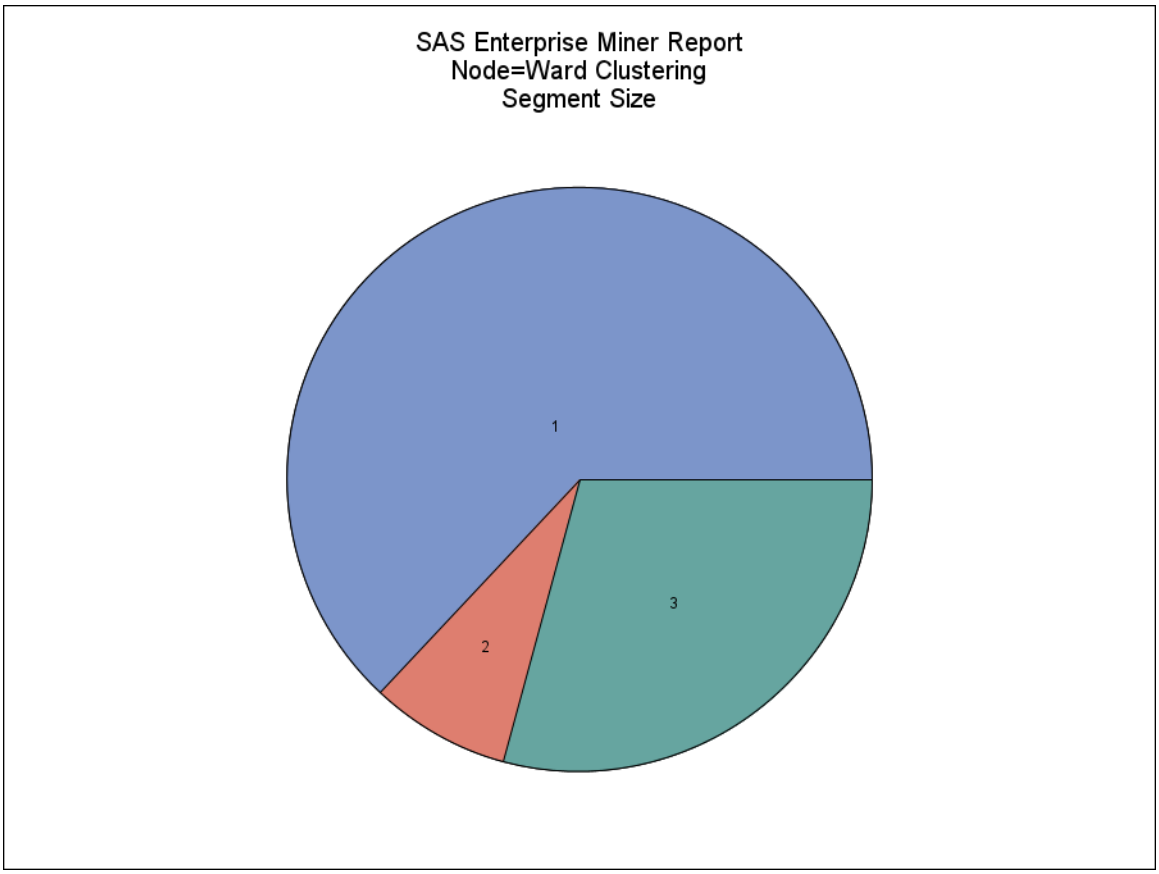
Node=Ward Clustering Variable Summary

Role	Level	Frequency Count	Name
INPUT	BINARY	8	AnyHealthcare CholCheck DiffWalk HeartDiseaseorAttack HighBP HighChol HvyAlcoholConsump NoDocbcCost
INPUT	INTERVAL	2	Age BMI
INPUT	ORDINAL	5	Education GenHlth Income MentHlth PhysHlth

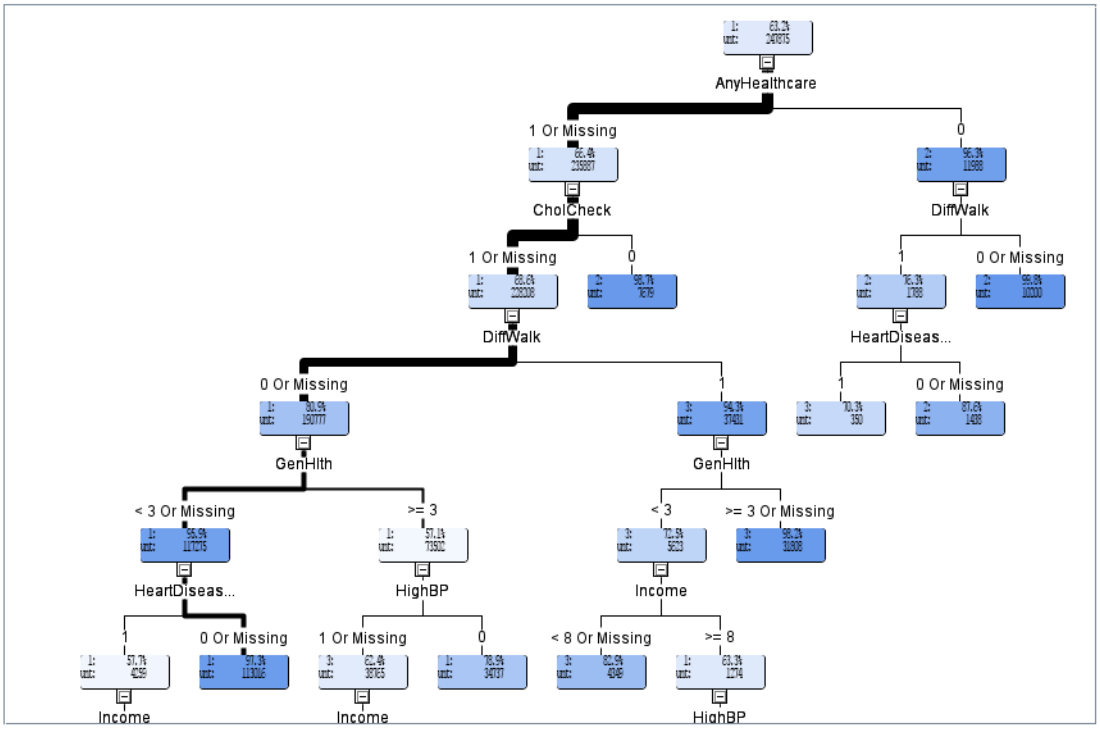
Node=Ward Clustering Mean Statistics

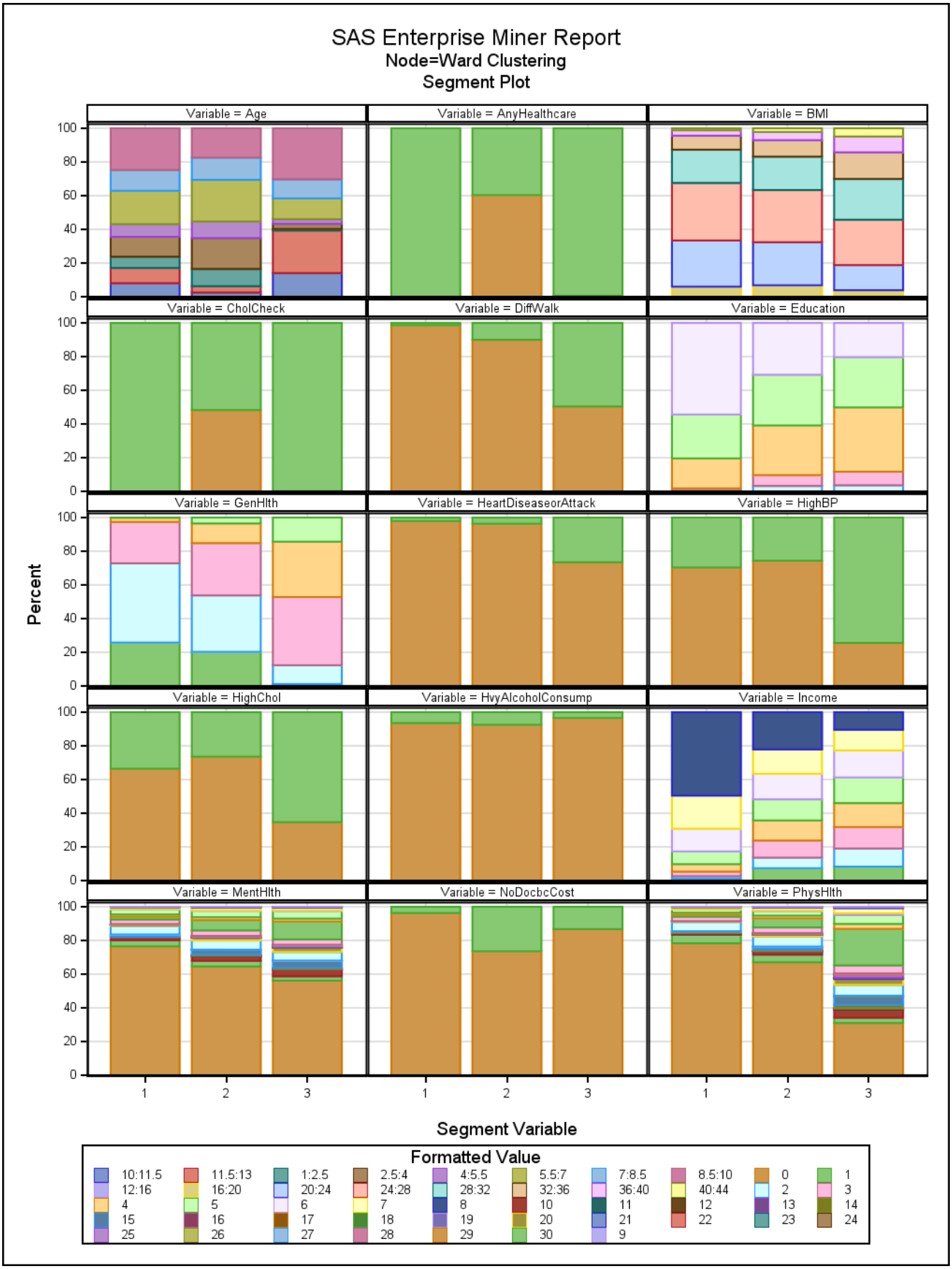
Variable	Segment 1	Segment 2	Segment 3
Clustering Criterion	0.72	0.72	0.72
Maximum Relative Change in Cluster Seeds	0.00	0.00	0.00
Improvement in Clustering Criterion	.	.	.
Frequency of Cluster	156554.00	19127.00	72194.00
Root-Mean-Square Standard Deviation	0.63	1.05	0.80
Maximum Distance from Cluster Seed	7.02	8.22	8.23
Nearest Cluster	3.00	1.00	1.00
Distance to Nearest Cluster	3.09	4.05	3.09
Age	7.55	6.34	9.59
BMI	26.98	27.47	29.75
AnyHealthcare=0	0.00	0.60	0.01
AnyHealthcare=1	1.00	0.40	0.99
CholCheck=0	0.00	0.48	0.00

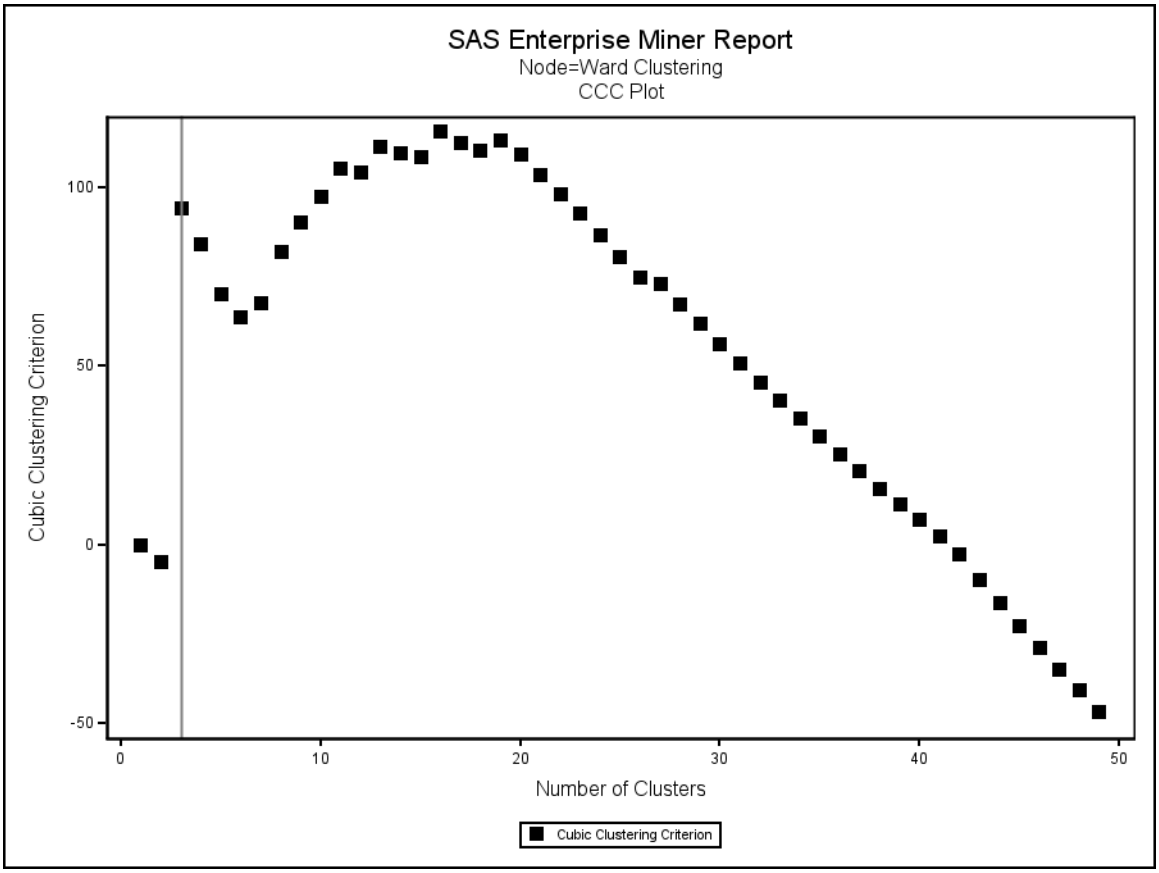
Variable	Segment 1	Segment 2	Segment 3
CholCheck=1	1.00	0.52	1.00
DiffWalk=0	0.99	0.90	0.50
DiffWalk=1	0.01	0.10	0.50
HeartDiseaseorAttack=0	0.98	0.96	0.73
HeartDiseaseorAttack=1	0.02	0.04	0.27
HighBP=0	0.70	0.74	0.26
HighBP=1	0.30	0.26	0.74
HighChol=0	0.66	0.74	0.35
HighChol=1	0.34	0.26	0.65
HvyAlcoholConsump=0	0.94	0.93	0.97
HvyAlcoholConsump=1	0.06	0.07	0.03
NoDocbcCost=0	0.96	0.74	0.87
NoDocbcCost=1	0.04	0.26	0.13
Transform: Education	0.57	0.43	0.36
Transform: GenHlth	0.38	0.49	0.75
Transform: Income	0.60	0.39	0.31
Transform: MentHlth	0.46	0.53	0.58
Transform: PhysHlth	0.41	0.48	0.69



SAS Enterprise Miner Report
Node=Ward Clustering
Tree Diagram




SAS Enterprise Miner Report
Node=Ward Clustering
Segment Plot
Variable = Age
Variable = AnyHealthcare
Variable = BMI
Variable = CholCheck
Variable = DiffWalk
Variable = Education
Variable = GenHlth
Variable = HeartDiseaseorAttack
Variable = HighBP
Variable = HighChol
Variable = HvyAlcoholConsump
Variable = Income
Variable = MentHlth
Variable = NoDocbcCost
Variable = PhysHlth
Percent
0
20
40
60
80
100
1
2
3
Segment Variable
Formatted Value
10:11.5
11.5:13
1:2.5
2.5:4
4:5.5
5.5:7
7:8.5
8.5:10
0
1
12:16
16:20
20:24
24:28
28:32
32:36
36:40
40:44
2
3
4
5
6
7
8
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
9



End of Report