```
 PRESERVE.
SET DECIMAL DOT.

GET DATA  /TYPE=TXT
  /FILE="\\kc.umkc.edu\kc-users\home\a\as9nb\My Documents\GP3\MLdata.csv"
  /ENCODING='UTF8'
  /DELIMITERS=","
  /QUALIFIER='"'
  /ARRANGEMENT=DELIMITED
  /FIRSTCASE=2
  /DATATYPEMIN PERCENTAGE=95.0
  /VARIABLES=
  case AUTO
  mode AUTO
  mean AUTO
  median AUTO
  std AUTO
  PC1 AUTO
  PC2 AUTO
  PC3 AUTO
  PC4 AUTO
  PC5 AUTO
  outcome AUTO
  /MAP.
RESTORE.
CACHE.
EXECUTE.

Data written to the working file.
11 variables and 200 cases written.
Variable: case              Type: Number  Format : F3
Variable: mode              Type: Number  Format : F3
Variable: mean              Type: Number  Format : F7.3
Variable: median            Type: Number  Format : F3
Variable: std               Type: Number  Format : F6.3
Variable: PC1               Type: Number  Format : E9.2
Variable: PC2               Type: Number  Format : F9.3
Variable: PC3               Type: Number  Format : F10.4
Variable: PC4               Type: Number  Format : F9.3
Variable: PC5               Type: Number  Format : F9.3
```

```
Variable: outcome            Type: Number  Format : F1

Substitute the following to build syntax for these data.
  /VARIABLES=
   case F3
   mode F3
   mean F7.3
   median F3
   std F6.3
   PC1 E9.2
   PC2 F9.3
   PC3 F10.4
   PC4 F9.3
   PC5 F9.3
   outcome F1

DATASET NAME DataSet3 WINDOW=FRONT.
DATASET ACTIVATE DataSet3.
DATASET CLOSE DataSet2.

SAVE OUTFILE='\\kc.umkc.edu\kc-users\home\a\as9nb\My Documents\GP3\CRTNNW.sav'
  /COMPRESSED.
* Decision Tree.
TREE outcome [n] BY mode [s] mean [s] median [s] std [s] PC1 [s] PC2 [s] PC3 [
s] PC4 [s] PC5 [s]
  /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES SCA
LE=AUTO
  /DEPCATEGORIES USEVALUES=[0 1] TARGET=[0 1]
  /PRINT MODELSUMMARY CLASSIFICATION RISK
  /GAIN CATEGORYTABLE=YES TYPE=[NODE] SORT=DESCENDING CUMULATIVE=NO
  /METHOD TYPE=CRT MAXSURROGATES=AUTO PRUNE=NONE
  /GROWTHLIMIT MAXDEPTH=AUTO MINPARENTSIZE=2 MINCHILDSIZE=1
  /VALIDATION TYPE=CROSSVALIDATION(10) OUTPUT=BOTHSAMPLES
  /CRT IMPURITY=GINI MINIMPROVEMENT=0.0001
  /COSTS EQUAL
  /PRIORS FROMDATA ADJUST=NO.
```

Classification Tree

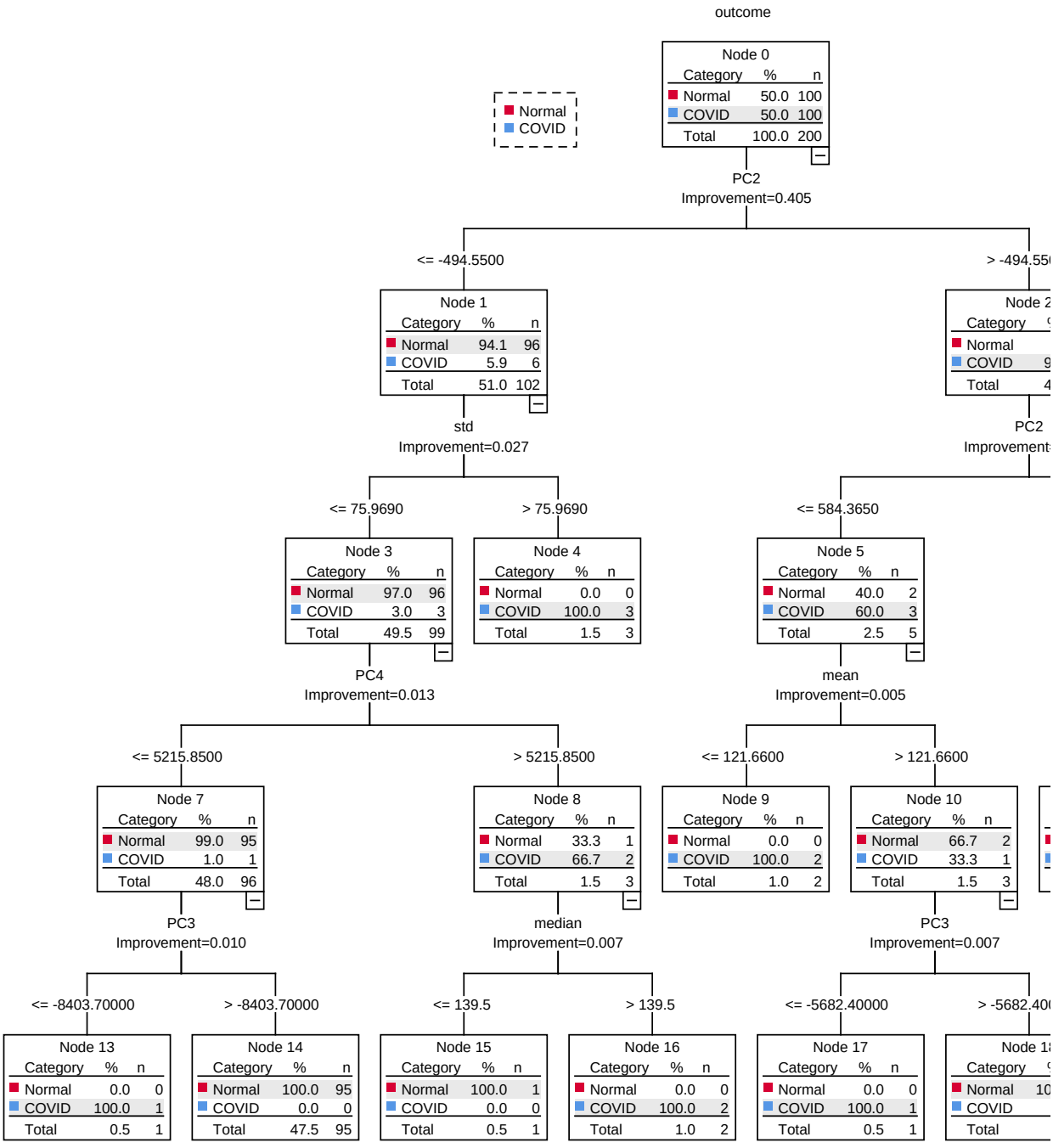
```
[DataSet3] \\kc.umkc.edu\kc-users\home\a\as9nb\My Documents\GP3\CRTNNW.sav
```

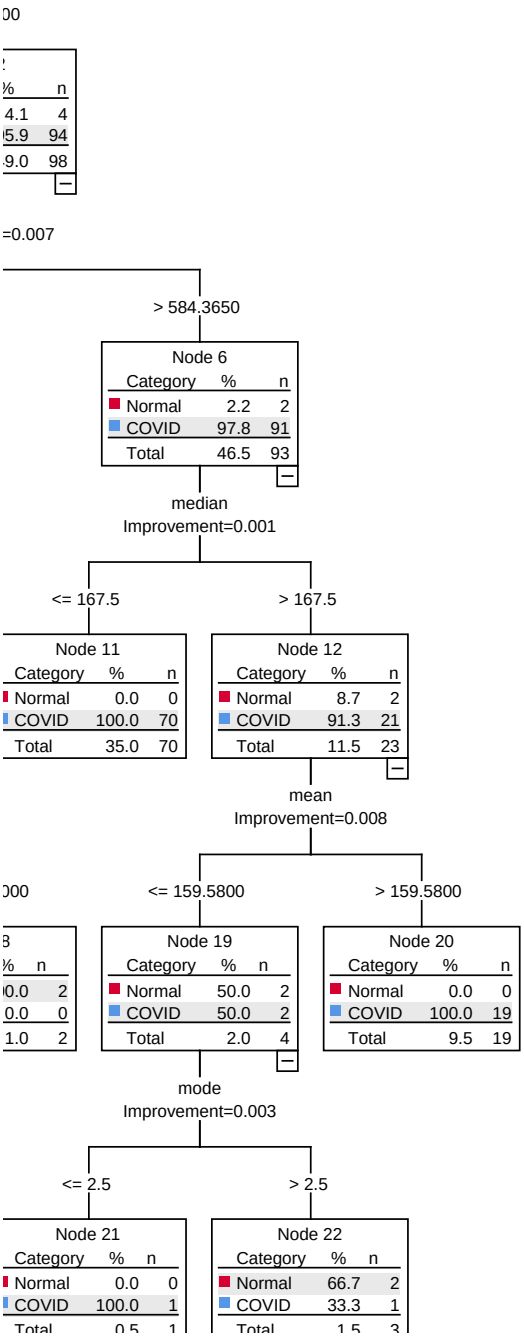
Warnings

Gain summary Tables are not displayed because profits are undefined.

Model Summary

Specifications	Growing Method	CRT
	Dependent Variable	outcome
	Independent Variables	mode, mean, median, std, PC1, PC2, PC3, PC4, PC5
	Validation	Cross Validation
	Maximum Tree Depth	5
	Minimum Cases in Parent Node	2
	Minimum Cases in Child Node	1
Results	Independent Variables Included	PC2, mean, median, mode, std, PC1, PC3, PC5, PC4
	Number of Nodes	23
	Number of Terminal Nodes	12
	Depth	5


outcome
Normal
COVID
Node 0
Category % n
Normal 50.0 100
COVID 50.0 100
Total 100.0 200
PC2
Improvement=0.405
<= -494.5500
> -494.55
Node 1
Category % n
Normal 94.1 96
COVID 5.9 6
Total 51.0 102
std
Improvement=0.027
Node 2
Category
Normal
COVID 9
Total 4
PC2
Improvement
<= 75.9690
> 75.9690
<= 584.3650
Node 3
Category % n
Normal 97.0 96
COVID 3.0 3
Total 49.5 99
Node 4
Category % n
Normal 0.0 0
COVID 100.0 3
Total 1.5 3
Node 5
Category % n
Normal 40.0 2
COVID 60.0 3
Total 2.5 5
PC4
Improvement=0.013
mean
Improvement=0.005
<= 5215.8500
> 5215.8500
<= 121.6600
> 121.6600
Node 7
Category % n
Normal 99.0 95
COVID 1.0 1
Total 48.0 96
Node 8
Category % n
Normal 33.3 1
COVID 66.7 2
Total 1.5 3
Node 9
Category % n
Normal 0.0 0
COVID 100.0 2
Total 1.0 2
Node 10
Category % n
Normal 66.7 2
COVID 33.3 1
Total 1.5 3
PC3
Improvement=0.010
median
Improvement=0.007
PC3
Improvement=0.007
<= -8403.70000
> -8403.70000
<= 139.5
> 139.5
<= -5682.40000
Node 13
Category % n
Normal 0.0 0
COVID 100.0 1
Total 0.5 1
Node 14
Category % n
Normal 100.0 95
COVID 0.0 0
Total 47.5 95
Node 15
Category % n
Normal 100.0 1
COVID 0.0 0
Total 0.5 1
Node 16
Category % n
Normal 0.0 0
COVID 100.0 2
Total 1.0 2
Node 17
Category % n
Normal 0.0 0
COVID 100.0 1
Total 0.5 1
Node 18
Category
Normal 10
COVID
Total


00
% n
4.1 4
5.9 94
9.0 98
=0.007
> 584.3650
Node 6
Category % n
Normal 2.2 2
COVID 97.8 91
Total 46.5 93
median
Improvement=0.001
<= 167.5
> 167.5
Node 11
Category % n
Normal 0.0 0
COVID 100.0 70
Total 35.0 70
Node 12
Category % n
Normal 8.7 2
COVID 91.3 21
Total 11.5 23
mean
Improvement=0.008
000
<= 159.5800
> 159.5800
% n
0.0 2
0.0 0
1.0 2
Node 19
Category % n
Normal 50.0 2
COVID 50.0 2
Total 2.0 4
Node 20
Category % n
Normal 0.0 0
COVID 100.0 19
Total 9.5 19
mode
Improvement=0.003
<= 2.5
> 2.5
Node 21
Category % n
Normal 0.0 0
COVID 100.0 1
Node 22
Category % n
Normal 66.7 2
COVID 33.3 1

Total	0.5	1

Total	1.5	3

Target Category: Normal

Gains for Nodes

Node	Node		Gain		Response	Index
	N	Percent	N	Percent		
14	95	47.5%	95	95.0%	100.0%	200.0%
18	2	1.0%	2	2.0%	100.0%	200.0%
15	1	0.5%	1	1.0%	100.0%	200.0%
22	3	1.5%	2	2.0%	66.7%	133.3%
11	70	35.0%	0	0.0%	0.0%	0.0%
20	19	9.5%	0	0.0%	0.0%	0.0%
4	3	1.5%	0	0.0%	0.0%	0.0%
16	2	1.0%	0	0.0%	0.0%	0.0%
9	2	1.0%	0	0.0%	0.0%	0.0%
13	1	0.5%	0	0.0%	0.0%	0.0%
17	1	0.5%	0	0.0%	0.0%	0.0%
21	1	0.5%	0	0.0%	0.0%	0.0%

Growing Method: CRT
Dependent Variable: outcome

Target Category: COVID

Gains for Nodes

	Node		Gain			
Node	N	Percent	N	Percent	Response	Index
11	70	35.0%	70	70.0%	100.0%	200.0%
20	19	9.5%	19	19.0%	100.0%	200.0%
4	3	1.5%	3	3.0%	100.0%	200.0%
16	2	1.0%	2	2.0%	100.0%	200.0%
9	2	1.0%	2	2.0%	100.0%	200.0%
13	1	0.5%	1	1.0%	100.0%	200.0%
17	1	0.5%	1	1.0%	100.0%	200.0%
21	1	0.5%	1	1.0%	100.0%	200.0%
22	3	1.5%	1	1.0%	33.3%	66.7%
14	95	47.5%	0	0.0%	0.0%	0.0%
18	2	1.0%	0	0.0%	0.0%	0.0%
15	1	0.5%	0	0.0%	0.0%	0.0%

Growing Method: CRT
Dependent Variable: outcome

Risk

Method	Estimate	Std. Error
Resubstitution	.005	.005
Cross-Validation	.080	.019

Growing Method: CRT
Dependent Variable: outcome

Classification

		Predicted	
Observed	Normal	COVID	Percent Correct
Normal	100	0	100.0%
COVID	1	99	99.0%
Overall Percentage	50.5%	49.5%	99.5%

Growing Method: CRT
Dependent Variable: outcome