```
Input ID X1 X2 X3 X4 X5 X6 X7 X8 X9 X10 X11 X12 X13 X14 X15 X16 X17 X18 X19
X20 X21 X22 X23;
*;
Data HBAT;
      Set HBAT;
      Label ID = 'ID - Identification Number'
              X1 = 'X1 - Customer Type'
              X2 = 'X2 - Industry Type'
              X3 = 'X3 - Firm Size'
              X4 = 'X4 - Region'
              X5 = 'X5 - Distribution System'
              X6 = 'X6 - Product Quality'
            X7 = 'X7 - E-Commerce'
          X8 = 'X8 - Technical Support'
          X9 = 'X9 - Complaint Resolution'
          X10 = 'X10 - Advertizing'
          X11 = 'X11 - Product Line'
          X12 = 'X12 - Salesforce Image'
          X13 = 'X13 - Competitive Pricing'
          X14 = 'X14 - Warranty & Claims'
          X15 = 'X15 - New Products'
          X16 = 'X16 - Order & Billing'
          X17 = 'X17 - Price Flexibility'
          X18 = 'X18 - Delivery Speed'
          X19 = 'X19 - Satisfaction'
          X20 = 'X20 - Likelihood of Recommendation'
          X21 = 'X21 - Likelihood of Future Purchase'
          X22 = 'X22 - Current Purchase/Usage Level'
          X23 = 'X23 - Consider Strategic Alliance/Partnership in Future';
*;
*;
* Select Variables ID X6 X8 X12 X15 X18;
*;
Data HBAT5;
      Set HBAT (Keep = ID X6 X8 X12 X15 X18);
*;
Proc Print Data = HBAT5;
*;
* Compute Variable Means;
*;
Proc Means Data = HBAT5;
      Var X6 X8 X12 X15 X18;
      Output Out = MeansHBAT
                  Mean(X6 X8 X12 X15 X18) = MeanX6 MeanX8 MeanX12 MeanX15
MeanX18;
*;
Proc Print Data = MeansHBAT;
*;
* Merge HBAT Data with HBAT Means ;
*;
Data MeansHBAT;
      Set MeansHBAT (Drop = _TYPE_ _FREQ_);
*;
Data HBATMeans;
      Retain ID X6 X8 X12 X15 X18;
      If _N_ = 1 Then Set MeansHBAT;
      Set HBAT;
```

```
*;
* Compute Centered HBAT Variables (Subtract Means) ;
*;
      X6C = X6 - MeanX6;
      X8C = X8 - MeanX8;
      X12C = X12 - MeanX12;
      X15C = X15 - MeanX15;
      X18C = X18 - MeanX18;
*;
* Compute Squared Centered HBAT Variables ;
*;
      X6CSQR = X6C ** 2;
      X8CSQR = X8C ** 2;
      X12CSQR = X12C ** 2;
      X15CSQR = X15C ** 2;
      X18CSQR = X18C ** 2;
*;
* Compute Totaled Squared Centered HBAT Variables ;
*;
      TotDiffSqr = Sum(X6CSQR,X8CSQR,X12CSQR,X15CSQR,X18CSQR);
*;
* Compute HBAT Variables Dissimalrities (Square Root of Total);
*;
      SqrRootTot = TotDiffSqr ** 0.5;
*;
* Rank the HBAT Variables Dissimalrities ;
*;
Proc Sort Data = HBATMeans;
      By Descending SqrRootTot;
*;
Proc Print Data = HBATMeans;
*;
****** Select 10 Largest HBAT Variables Dissimalrities *****;
*;
Data HBATMeans10;
      Set HBATMeans (Keep = X6C X8C X12C X15C X18C X6CSQR X8CSQR X12CSQR
X15CSQR X18CSQR SqrRootToT);
      If _N_ LE 10;
*;
Proc Print Data = HBATMeans10;
*;
* SAS Hierarchical Cluster Analysis ;
*;
* The PROC CLUSTER statement starts the CLUSTER procedure, specifies a
clustering method, and
      optionally specifies details for clustering methods, data sets, data
processing, and displayed output.
*;
* The METHOD = specification determines the clustering method used by the
procedure. Any one of
      the 11 methods can be specified for name;
*;
*     WARD | WAR requests Ward's minimum-variance method (error sum of
squares, trace W).
           Distance data are squared unless you specify the NOSQUARE option.
           To reduce distortion by outliers, the TRIM= option is recommended.
           See the NONORM option.;
```

```
*;
*     NONORM - prevents the distances from being normalized to unit mean or
unit root mean square with
               most methods. With METHOD=WARD, the NONORM option prevents the
between-cluster
               sum of squares from being normalized by the total sum of
squares to yield a squared semipartial
               correlation;
*;
*     SIMPLE | S - displays means, standard deviations, skewness, kurtosis,
and a coefficient of bimodality. The
                    SIMPLE option applies only to coordinate data.;
*;
*     CCC - displays the cubic clustering criterion and approximate expected
R square under the uniform
            null hypothesis. The statistics associated with the RSQUARE
option, R square
            and semipartial R square, are also displayed. The CCC option
applies only to coordinate
            data. The CCC option is not appropriate with METHOD=SINGLE
because of the method's
            tendency to chop off tails of distributions. Computation of the
CCC requires the eigenvalues
            of the covariance matrix. If the number of variables is large,
computing the eigenvalues
            requires much computer time and memory.;
*;
*     PSEUDO - displays pseudo F and t2 statistics. This option is effective
only when the data are coordinates
               or when METHOD=AVERAGE, METHOD=CENTROID, or METHOD=WARD is
specified.;
*;
*     RMSSTD - displays the root mean square standard deviation of each
cluster. This option is effective only
               when the data are coordinates or when METHOD=AVERAGE,
METHOD=CENTROID, or
               METHOD=WARD is specified.;
*;
*     RSQUARE | RSQ - displays the R square and semipartial R square. This
option is effective only when the data
                      are coordinates or when METHOD=AVERAGE or
METHOD=CENTROID is specified. The
                      R square and semipartial R square statistics are always
displayed with METHOD=WARD.;
*;
*     OUTTREE = SAS-data-set
              - creates an output data set that can be used by the TREE
procedure to draw a tree diagram. You
                must give the data set a two-level name to save it. See SAS
Language Reference: Concepts
                for a discussion of permanent data sets. If you omit the
OUTTREE= option, the data set is
                named by using the DATAn convention and is not permanently
saved. If you do not want to
                create an output data set, use OUTTREE=_NULL_.;
*;
*;
```

```
Proc Cluster Data=HBATMeans Method=Ward NoNorm Simple CCC Pseudo RmsStd
RSquare OutTree=Tree;
      Var X6 X8 X12 X15 X18;
*;
* Plot the Dendrogram;
*;
Proc Tree Data=Tree;
*;
* Remove Identified Outliers: Observations 6 and 87;
*;
Data HBATMeans98Obs;
      Set HBATMeans;
      If ID EQ 6 Then Delete;
      If ID EQ 87 Then Delete;
*;
***** SAS Hierarchical Cluster Analysis *****;
*;
Proc Cluster Data=HBATMeans98Obs Method=Ward NoNorm Simple CCC Pseudo RmsStd
RSquare OutTree=Tree;
      Var X6 X8 X12 X15 X18;
*;
* Plot the Dendrogram;
*;
Proc Tree Data=Tree;
*;
*;
*;
***** SAS Non-Hierarchical 4-Cluster Analysis *****;
*;
* The FASTCLUS procedure performs a disjoint cluster analysis on the basis of
distances computed
               from one or more quantitative variables. The observations are
divided into clusters such that every
               observation belongs to one and only one cluster, the clusters
do not form a tree structure as they do
               in the CLUSTER procedure.;
*;
* The FASTCLUS procedure combines an effective method for finding initial
clusters with a standard
               iterative algorithm for minimizing the sum of squared
distances from the cluster means.
               The result is an efficient procedure for disjoint clustering
of large data sets.;
*;
*     RADIUS = t R=t
             - establishes the minimum distance criterion for selecting new
seeds. No observation is considered
             as a new seed unless its minimum distance to previous seeds
exceeds the value given
             by the RADIUS= option. The default value is 0. If you specify
the REPLACE=RANDOM
             option, the RADIUS= option is ignored.;
*;
*     RANDOM = n
          - specifies a positive integer as a starting value for the pseudo-
random number generator for
```

```
            use with REPLACE=RANDOM. If you do not specify the RANDOM=
option, the time of
            day is used to initialize the pseudo-random number sequence.
            REPLACE = FULL | PART | NONE | RANDOM
                              specifies how seed replacement is performed, as
follows:
                        FULL requests default seed replacement
                        PART requests seed replacement only when the distance
between the observation
                              and the closest seed is greater than the minimum
distance between seeds.
                       NONE suppresses seed replacement.
                       RANDOM selects a simple pseudo-random sample of
complete observations as initial
                                cluster seeds.;
*;
*     MAXCLUSTERS = n MAXC = n
               - specifies the maximum number of clusters permitted. If you
omit the MAXCLUSTERS=
                 option, a value of 100 is assumed.;
*;
*     MAXITER = n
          - specifies the maximum number of iterations for recomputing
cluster seeds.;
*;
*     LIST - lists all observations, giving the value of the ID variable (if
any), the number of the cluster
          to which the observation is assigned, and the distance between the
observation and the final
          cluster seed.;
*;
*     DISTANCE | DIST - computes distances between the cluster means.;
*;
*     OUT = SAS-data-set
          - creates an output data set to contain all the original data, plus
the new variables CLUSTER and
            DISTANCE.;
*;
Proc FastClus Data=HBATMeans98Obs Radius=0 Replace=Random MaxClusters=4
Maxiter=20 List Distance Out=Clust;
      Var X6 X8 X12 X15 X18;
*;
* Plot 4-Cluster Obs Membership with X-Y Variable Scatterplots ;
*;
Proc Print Data = Clust;
*;
Proc Sgplot Data = Clust;
      Scatter X = X6 Y = X8 / Group=Cluster ;
*;
Proc Sgplot Data = Clust;
      Scatter X = X6 Y = X12 / Group=Cluster ;
*;
Proc Sgplot Data = Clust;
      Scatter X = X6 Y = X15 / Group=Cluster ;
*;
Proc Sgplot Data = Clust;
      Scatter X = X6 Y = X18 / Group=Cluster ;
```

```
*;
Proc Sgplot Data = Clust;
      Scatter X = X8 Y = X12 / Group=Cluster ;
*;
Proc Sgplot Data = Clust;
      Scatter X = X8 Y = X15 / Group=Cluster ;
*;
Proc Sgplot Data = Clust;
      Scatter X = X8 Y = X18 / Group=Cluster ;
*;
Proc Sgplot Data = Clust;
      Scatter X = X12 Y = X15 / Group=Cluster ;
*;
Proc Sgplot Data = Clust;
      Scatter X = X12 Y = X18 / Group=Cluster ;
*;
Proc Sgplot Data = Clust;
      Scatter X = X15 Y = X18 / Group=Cluster ;
*;
***** Validation and Profiling the 4-Clusters *****;
*;
* Merge Cluster Assignments with Original HBAT Data By ID (with Outliers
Removed);
*;
* Select Variables ID X15 X20 X21 X22;
*;
Data HBAT4;
      Set HBAT (Keep = ID X15 X20 X21 X22);
      If ID EQ 6 Then Delete;
      If ID EQ 87 Then Delete;
*;
Proc Sort Data = HBAT4;
      By ID;
Proc Sort Data = Clust;
      By ID;
Data HBAT4Clust (Keep = ID X15 X20 X21 X22 Cluster);
      Merge HBAT4 Clust;
      By ID;
Proc Print Data = HBAT4Clust;
*;
***** Assessing 4-Cluster Criterion Validity *****;
*;
* GLM MANOVA Analysis ;
*;
Proc GLM Data = HBAT4Clust;
    Class Cluster;
      Model X15 X20 X21 X22 = Cluster;
      Means Cluster / Scheffe Tukey LSD SNK Duncan;
      Means Cluster / Hovtest = Levene Hovtest = bf Hovtest = Bartlett;
      Means Cluster;
      Manova H = Cluster / MStat = Exact;
*;
***** Profiling the Final 4-Cluster Solution *****;
*;
* Merge Cluster Assignments with Original HBAT Data By ID (with Outliers
Removed);
*;
```

```
* Select Variables ID X1 X2 X3 X4 X5;
*;
Data HBAT5;
      Set HBAT (Keep = ID X1 X2 X3 X4 X5);
      If ID EQ 6 Then Delete;
      If ID EQ 87 Then Delete;
*;
Proc Sort Data = HBAT4;
      By ID;
Proc Sort Data = Clust;
      By ID;
Data HBAT5Clust (Keep = ID X1 X2 X3 X4 X5 Cluster);
      Merge HBAT5 Clust;
      By ID;
Proc Print Data = HBAT5Clust;
*;
***** Cross-Classification of Clusters on X1 X2 X3 X4 X5 *****;
*;
Proc Freq Data = HBAT5Clust;
      Table Cluster * X1;
*;
Proc Freq Data = HBAT5Clust;
      Table Cluster * X2;
*;
Proc Freq Data = HBAT5Clust;
      Table Cluster * X3;
*;
Proc Freq Data = HBAT5Clust;
      Table Cluster * X4;
*;
Proc Freq Data = HBAT5Clust;
      Table Cluster * X5;
*;
*;
*;
***** SAS Non-Hierarchical 3-Cluster Analysis *****;
*;
Proc FastClus Data=HBATMeans98Obs Radius=0 Replace=Random MaxClusters=3
Maxiter=20 List Distance Out=Clust;
      Var X6 X8 X12 X15 X18;
*;
* Plot 3-Cluster Obs Membership with X-Y Variable Scatterplots ;
*;
Proc Print Data = Clust;
*;
Proc Sgplot Data = Clust;
      Scatter X = X6 Y = X8 / Group=Cluster ;
*;
Proc Sgplot Data = Clust;
      Scatter X = X6 Y = X12 / Group=Cluster ;
*;
Proc Sgplot Data = Clust;
      Scatter X = X6 Y = X15 / Group=Cluster ;
*;
Proc Sgplot Data = Clust;
      Scatter X = X6 Y = X18 / Group=Cluster ;
*;
```

```
Proc Sgplot Data = Clust;
      Scatter X = X8 Y = X12 / Group=Cluster ;
*;
Proc Sgplot Data = Clust;
      Scatter X = X8 Y = X15 / Group=Cluster ;
*;
Proc Sgplot Data = Clust;
      Scatter X = X8 Y = X18 / Group=Cluster ;
*;
Proc Sgplot Data = Clust;
      Scatter X = X12 Y = X15 / Group=Cluster ;
*;
Proc Sgplot Data = Clust;
      Scatter X = X12 Y = X18 / Group=Cluster ;
*;
Proc Sgplot Data = Clust;
      Scatter X = X15 Y = X18 / Group=Cluster ;
*;
***** Validation and Profiling the 3-Clusters *****;
*;
* Merge Cluster Assignments with Original HBAT Data By ID (with Outliers
Removed);
*;
* Select Variables ID X15 X20 X21 X22;
*;
Data HBAT4;
      Set HBAT (Keep = ID X15 X20 X21 X22);
      If ID EQ 6 Then Delete;
      If ID EQ 87 Then Delete;
*;
Proc Sort Data = HBAT4;
      By ID;
Proc Sort Data = Clust;
      By ID;
Data HBAT4Clust (Keep = ID X15 X20 X21 X22 Cluster);
      Merge HBAT4 Clust;
      By ID;
Proc Print Data = HBAT4Clust;
*;
***** Assessing 3-Cluster Criterion Validity *****;
*;
* GLM MANOVA Analysis ;
*;
Proc GLM Data = HBAT4Clust;
    Class Cluster;
      Model X15 X20 X21 X22 = Cluster;
      Means Cluster / Scheffe Tukey LSD SNK Duncan;
      Means Cluster / Hovtest = Levene Hovtest = bf Hovtest = Bartlett;
      Means Cluster;
      Manova H = Cluster / MStat = Exact;
*;
***** Profiling the Final 3-Cluster Solution *****;
*;
* Merge Cluster Assignments with Original HBAT Data By ID (with Outliers
Removed);
*;
* Select Variables ID X1 X2 X3 X4 X5;
```

```
*;
Data HBAT5;
      Set HBAT (Keep = ID X1 X2 X3 X4 X5);
      If ID EQ 6 Then Delete;
      If ID EQ 87 Then Delete;
*;
Proc Sort Data = HBAT4;
      By ID;
Proc Sort Data = Clust;
      By ID;
Data HBAT5Clust (Keep = ID X1 X2 X3 X4 X5 Cluster);
      Merge HBAT5 Clust;
      By ID;
Proc Print Data = HBAT5Clust;
*;
***** Cross-Classification of Clusters on X1 X2 X3 X4 X5 *****;
*;
Proc Freq Data = HBAT5Clust;
      Table Cluster * X1;
*;
Proc Freq Data = HBAT5Clust;
      Table Cluster * X2;
*;
Proc Freq Data = HBAT5Clust;
      Table Cluster * X3;
*;
Proc Freq Data = HBAT5Clust;
      Table Cluster * X4;
*;
Proc Freq Data = HBAT5Clust;
      Table Cluster * X5;
*;
*;
*;
*     ods graphics off;
*;
*;
Run;
Quit;
```