

*SegmentSolve*TM

QuickStart Manual

For the SegmentSolveTM Segmentation Decision Support Software



Decision Support Sciences. Better Science. Better Solutions.

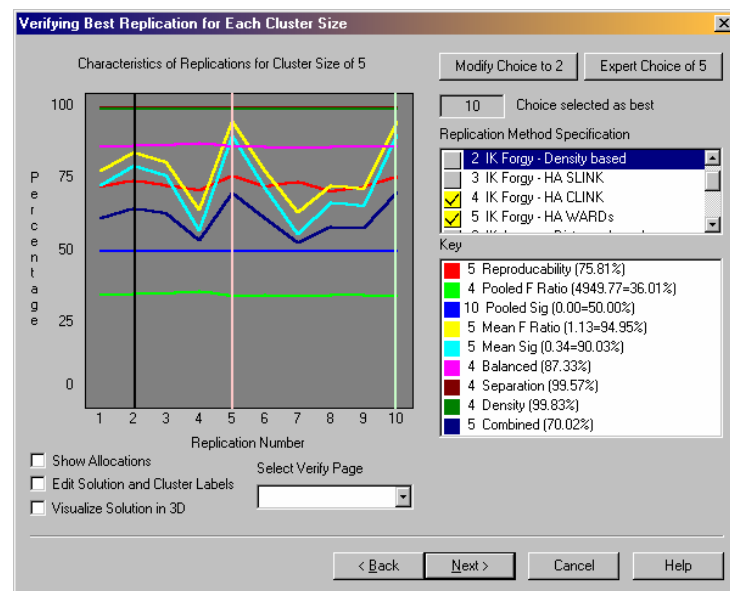
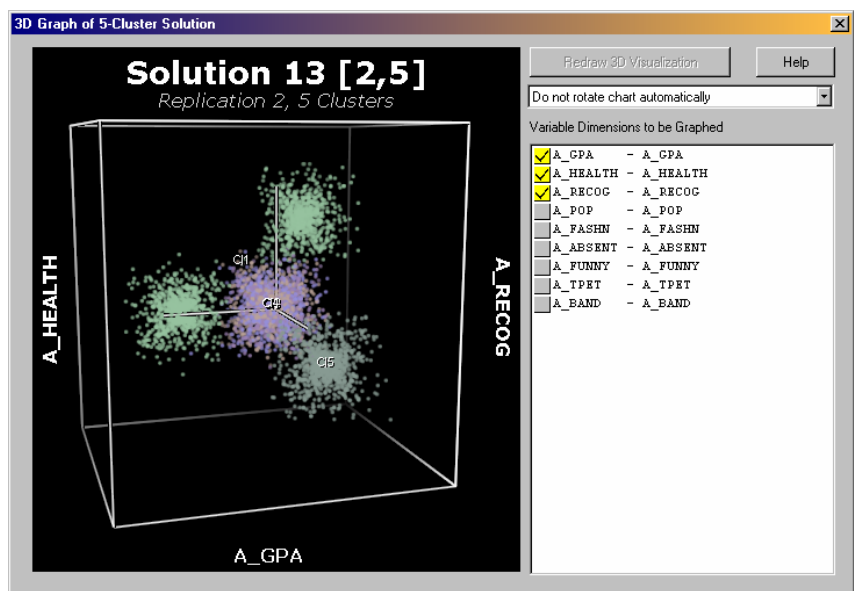
Decision Support Sciences, Inc., East West Technology Center
1751 W. Diehl Road, Suite 150, Naperville, IL 60563
www.DecisionSupportSciences.com
(630) 428-1847



What is Market Segmentation or Segment Analysis?

Segmentation is the process of classifying a population into segments based on similarities in the underlying data. This data comes from either a survey or from an existing data mart or database. Segmentation, as usually deployed, uses a mathematical process to analyze the data to find groups of customers or prospects who have similar values or similar behaviors. (But, the things that you are trying to group don't have to be people, they can be anything that can be described by a set of variables.)

SegmentSolve™ uses 13 different forms of cluster analysis to perform the segmentation. But SegmentSolve™ is more than a multi-algorithm cluster analysis engine. SegmentSolve™ is an end-to-end segmentation engine that provides segmentation decision support. As such, it is Decision Support Software (DSS). As a DSS tool, it gets you to the *solution* - it visualizes and evaluates the resultant clusters to determine the best overall solution out of hundreds (or thousands) tried.



Installing SegmentSolve™

System Requirements:

- SegmentSolve™ will only run with IBM-compatible PC's that are Pentium class, running Windows 98 or higher.
- It is recommended that you have at least 20MB of hard drive space, and 64 MB RAM.

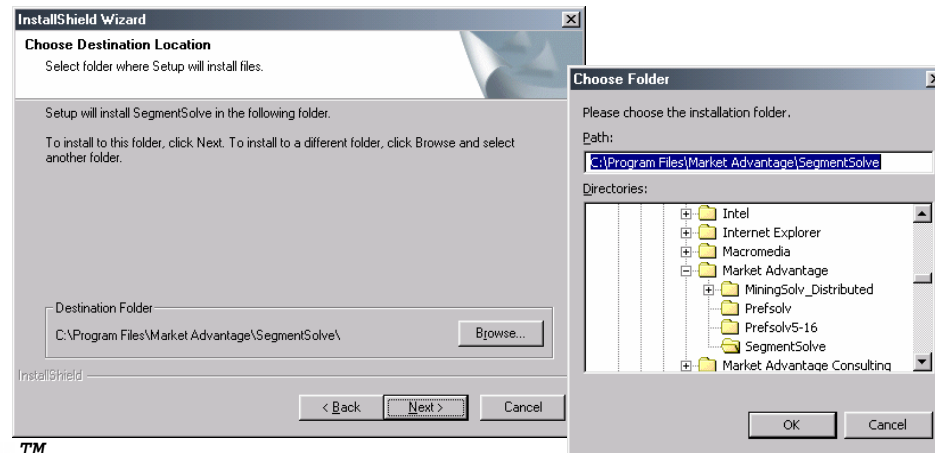
Installation Steps:

If you have the installation CD:

Insert the CD in your CD-ROM drive. If the installation program does not run automatically, browse the CD and double-click the file Setup.exe.

If you downloaded the installation file:

Browse to the directory where it's been saved and double-click the file Setup.exe. The installer guides you through the installation process, allowing you to select the directory where program files will be stored.

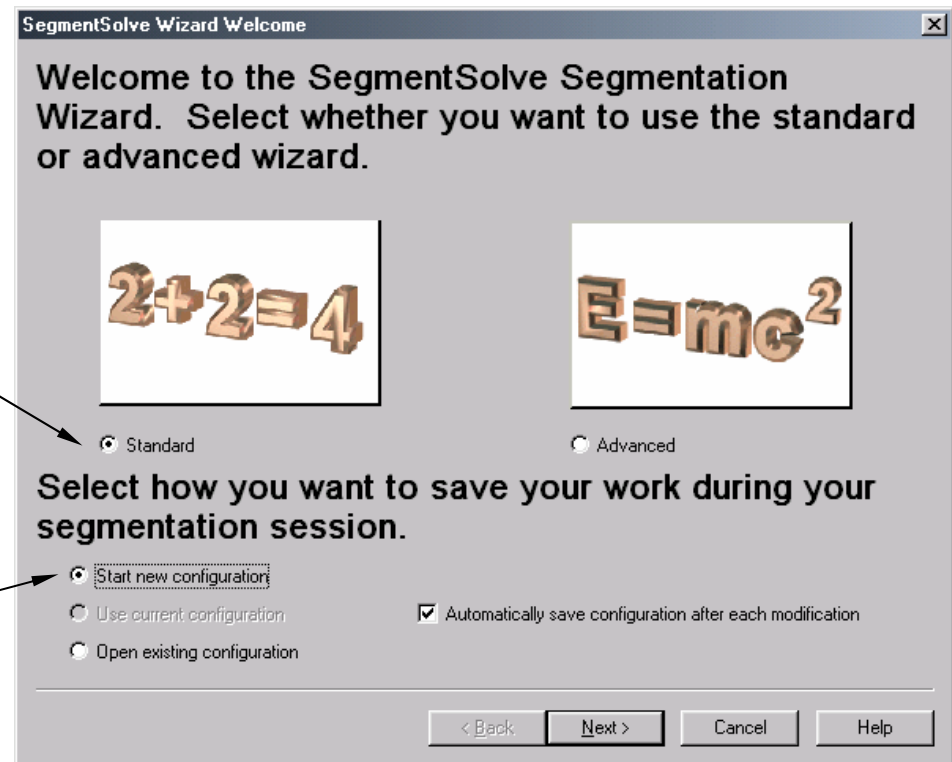


SegmentSolve™ Wizard Tutorial

SegmentSolve™ Wizard Step 1

■ Introduction

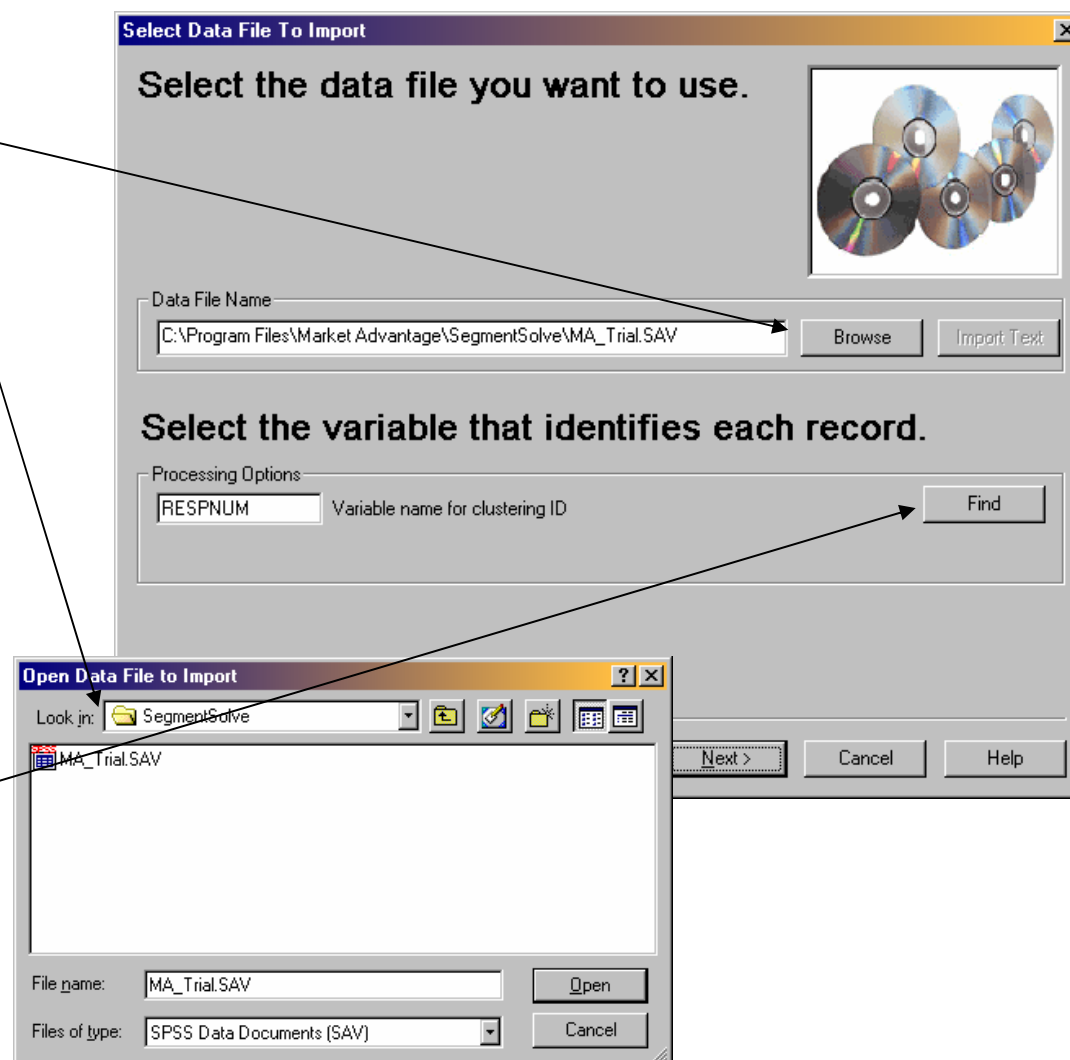
- This tutorial will demonstrate how to perform a segmentation with SegmentSolve™ using the data file provided, MA_Trial.SAV.
- The wizard can be used at two different levels. The standard settings are used if you don't want to modify the default statistical settings for the cluster analysis. This still allows you to customize the variables and filters used in the analysis. For this example, the standard wizard will be used. (Refer to the SegmentSolve™ manual for descriptions of options available at the advanced level.)
- As you go through the wizard, all of the settings you choose can be saved in a SegmentSolve™ configuration file. Select "Start new configuration" to begin a new configuration file. By default, SegmentSolve™ does not save the configuration file after each step of the wizard. However, this default can be overridden by checking the check box to the right.
- Click the Next button, and you will be prompted to type in a name for the configuration file.



SegmentSolve™ Wizard Step 2

■ Select the Data File

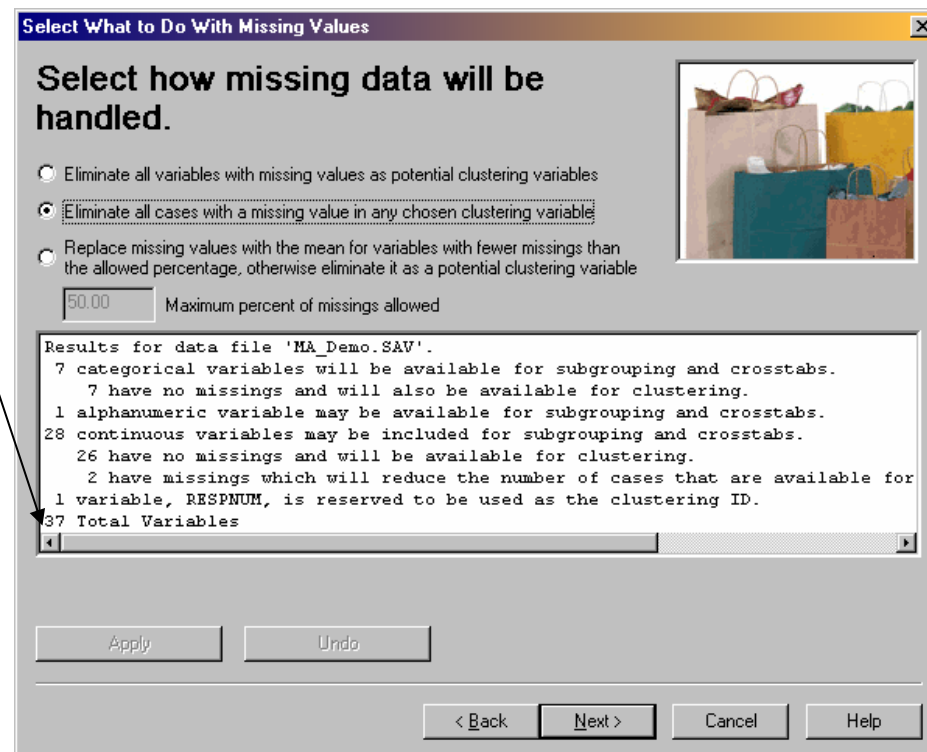
- To select the data file, click on the Browse button.
- The default directory is the SegmentSolve™ installation directory which should contain the sample data file MA_Demo.SAV. Highlight the file and select Open. The file name and path will then appear in the Data File Name box. If you are using the demo version of SegmentSolve™, this is the only file that SegmentSolve™ will allow you to use.
- A clustering ID must be specified to run the cluster analysis. A clustering ID is a variable from the data file that has a unique value for each record in the file.
- Select the Find button to display all the variables. Highlight the variable RESPNUM and select the Use button.
- Click the Next button on the wizard to go to the next step.



SegmentSolve™ Wizard Step 3

■ Select What to Do With Missing Values

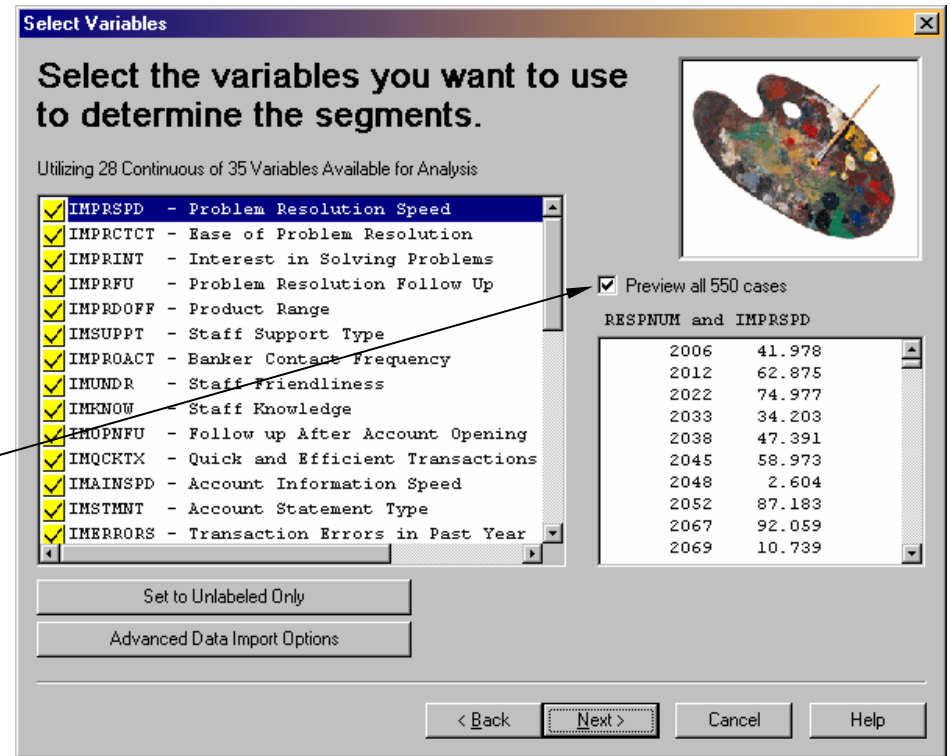
- Variables that have missing data for some cases can not be used in the cluster analysis. If there are any missing values, SegmentSolve™ gives three options for handling them. Only one option may be selected at one time.
- For this example, use the second option, which is the default. All cases (respondents) that have a missing value for any variable will not be included in the cluster analysis. In the sample data file, there are 24 cases with missing values that will be excluded.
- In the large text box, the number of categorical and continuous variables included in or excluded from the analysis with the current maximum percentage of missings will be displayed. The categorical variables are those that are used for filtering or selecting portions of the data. The continuous variables are those that do not have level labels, and are the default variables that will be used for clustering.
- Click the Next button on the wizard to go to the next step.



SegmentSolve™ Wizard Step 3

■ Clustering Variables

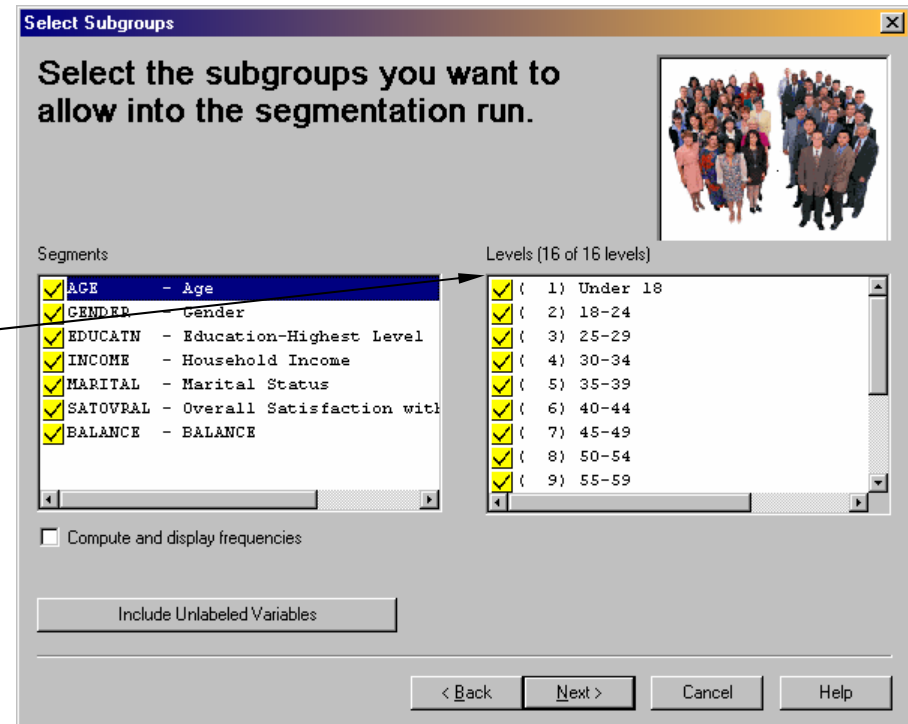
- The continuous variables available for the analysis are displayed on this dialog. Variables with a checkmark to the left will be included, and variables that are not checked will be excluded from the analysis. By default, all available variables are selected. Use the mouse to click on the checkmark if you want to exclude a variable from the analysis.
- For this tutorial we will keep all of the variables selected.
- Select the 'Preview Variables' option to view the values for any of the variables. Highlight the variable in the left box to view the values in the Preview box on the right. The clustering ID is listed in the first column, and the cluster variable in the second column.
- Click the Next button to continue to the next wizard step.



SegmentSolve™ Wizard Step 4

■ Select Subgroups

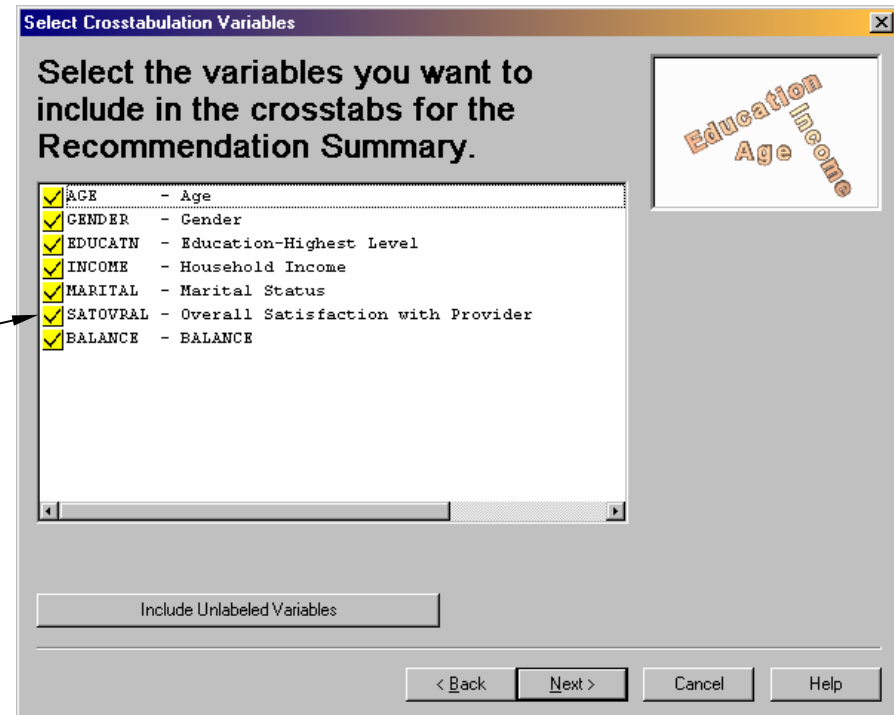
- All categorical variables from the data file will be shown in the left box of this dialog. Specific levels of each of the variables may be selected to include or exclude certain “cases” from the analysis. (A case is the individual set of data that you are trying to group. In marketing this is usually the person who took the survey.)
- The “Segments” box on the left displays the variable names and labels.
- The “Levels” box on the right displays the levels of the highlighted variable. By default, all levels are selected. To exclude one level from the analysis, click on the yellow box to deselect it. In this example, we have left all subgroup variables selected.
- Select the Next button at the bottom of the dialog to move on to the next step of the clustering wizard.



SegmentSolve™ Wizard Step 5

■ Select Crosstabulation Variables

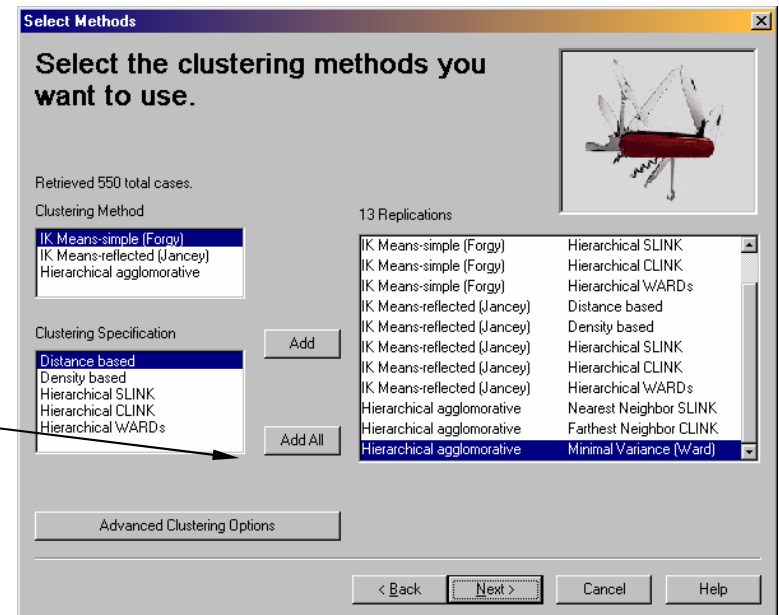
- The recommendation summary is a report created after clustering that characterizes the clusters. These cluster characterizations are based on the subgroup variables you select in this step.
- All of the categorical variables available for the crosstabs are shown in the list. By default, all of them are selected. Click on the checkbox to the left of a variable to deselect it.
- Again, in this example we have left all of the variables selected. Click the Next button to go on.



SegmentSolve™ Wizard Step 6

■ Select Clustering Methods

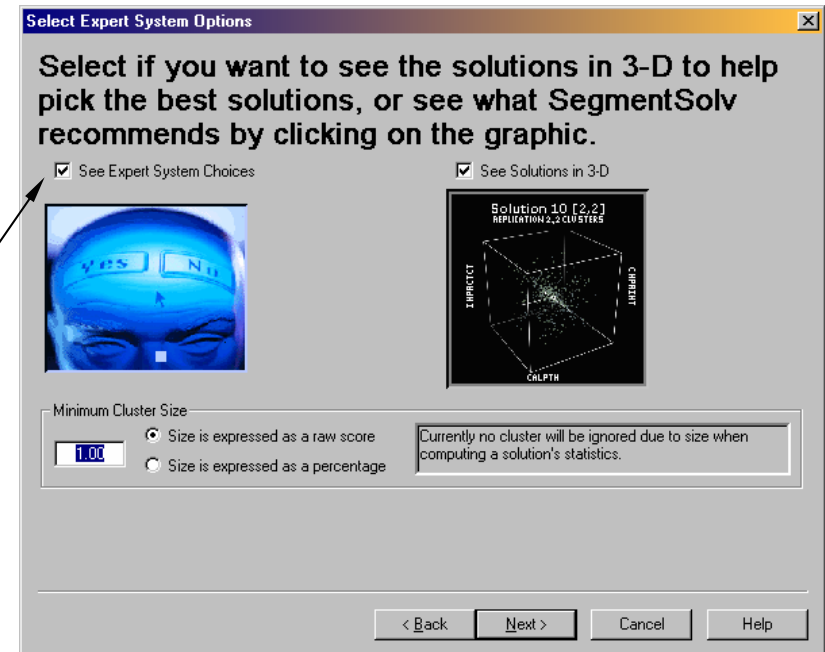
- The clustering methods used for the analysis are selected at this step. The combination of a clustering method and how that cluster method groups cases is called a replication. There are 13 possible replications. By default, all replications are included.
- A replication can be included multiple times. (This is because many of the replications use an algorithm that starts with a semi-random starting solution. So running the replication many times does not produce the same final solution.)
- To add a single replication, highlight a method and specification in the left boxes on the dialog and click the **Add** button. To add all of the possible replications, use the **Add All** button.
- To remove a replication from the selected list, use the mouse to highlight the replication in the list at right. The **Add** button becomes the **Remove** button. Click the **Remove** button to eliminate the replication.
- For this example, leave the replication list the way it is. All 13 replications should be selected. By default the 2-10 cluster solutions will be developed. (To change this click on the Advanced Clustering Options button.) Thus, 117 total cluster analyses will be run (13 replications X 9 cluster solutions).
- Select the **Next** button, and SegmentSolve™ will automatically begin calculating the clusters.
- While SegmentSolve™ performs the cluster analysis, it provides visual feedback. The circles represent the proportional sizes of the clusters. At the top of the cluster visualization dialog are the details for the current analysis being run.



SegmentSolve™ Wizard Step 7

■ Select Expert System Options

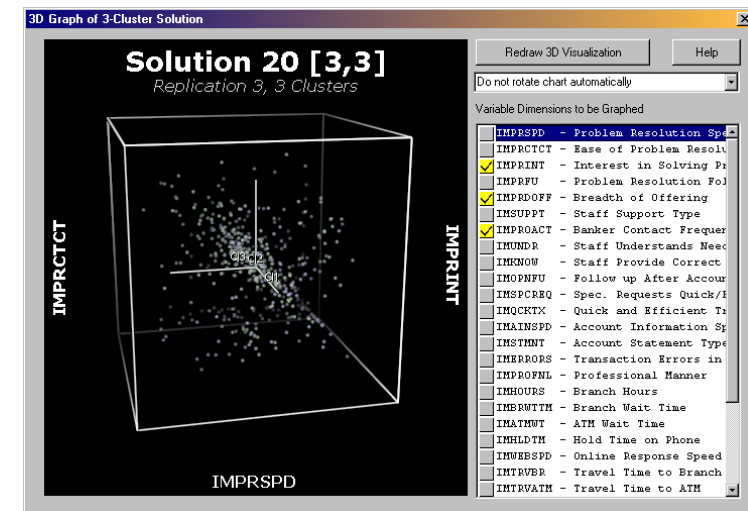
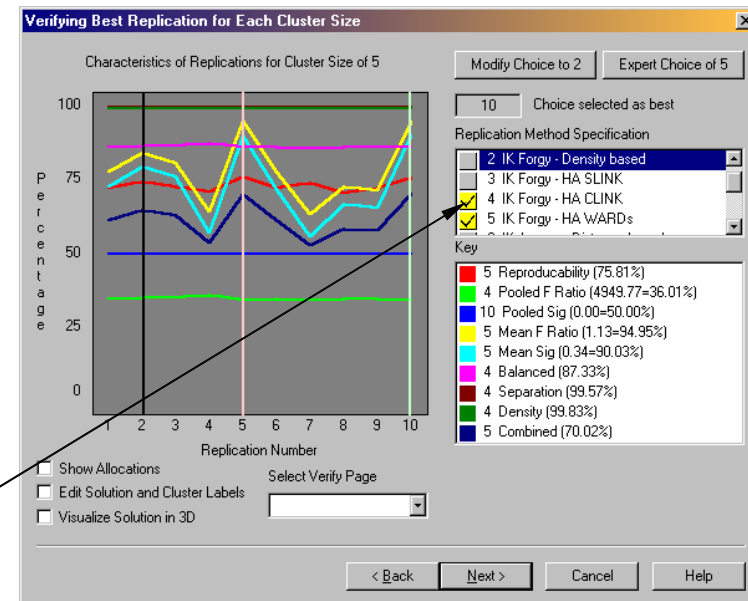
- Of all the solutions generated, SegmentSolve™ chooses one from each cluster size as the “expert system choice” for that cluster size, based on multiple mathematical criteria.
- To inspect the all the clusters that were created, check the option on the left to **See Expert System Choices**. This option allows you to modify the expert system choices before any reports are generated. If this option is not checked, SegmentSolve™ will skip to the processing step where reports are created, retaining the single best solution identified for each cluster size.
- The **See Solutions in 3D** option will bring up the Visualize in 3D dialog in the next wizard step. This option is checked by default; use the mouse to uncheck the option if you do not want to view the cluster solution visualization.
- The **minimum cluster size** refers to the minimum number of cases that can be in a cluster in order for SegmentSolve™ to include it in calculations for statistical analysis (e.g. pooled F ratio or mean significance). If **percentage** is selected, then each cluster must contain at least the specified minimum percentage of respondents.
- Click on the **Next** button to move forward in the wizard. A progress bar will appear while the expert system choices are determined.



SegmentSolve™ Wizard Step 8

■ Verify Best Replications

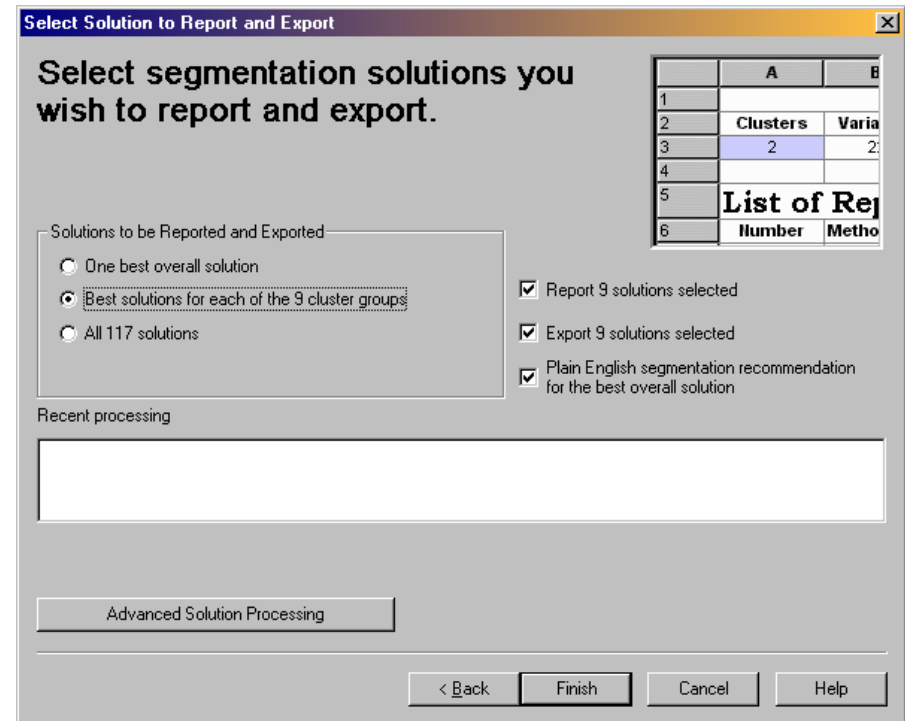
- If the option to See **Expert System Choices** was selected in the previous wizard step, this dialog will display how each of the cluster solutions performs based on 9 different statistical criteria. One screen is shown for each cluster size. The example to the right shows the results for the 5 cluster solution.
- One solution for each cluster size is selected automatically as the expert system choice. In this case it is replication #5 as reported on the Expert Choice button in the upper right of the wizard dialog box. To select additional replications to be reported or exported, click on the checkbox next to the replication name in the replication list. All replications marked with a checkmark will be stored for optional reporting and exporting.
- Click **Visualize Solution in 3D** to show or hide the 3D graph of the clusters. The graph shows only the selected replication. Because the graph is 3-dimensional, only three variables can be graphed at once. Use the mouse to select or de-select which variables you want to display, then select the **Redraw Visualization** button.
- Click on the **Next** button to proceed through the verification dialogs.



SegmentSolve™ Wizard Step 9

■ Report and Export Solutions

- Reports can be created to show how the clusters are differentiated and give the means, deviations, and other statistics for each cluster group. By default, these reports are created in spreadsheet format. One report gives the detailed statistical information for the clusters, and the other report is a recommendation summary that characterizes the clusters based on the clustering and demographic variables provided.
- Solutions are also **exported** to an ASCII data file using the clustering ID specified earlier in the wizard.
- For this example, leave the defaults set so that all three report/export options are checked on the right side, and “Best solutions for each of the 9 cluster groups” is selected on the left.
- Click on the **Finish** button to generate the report and export files and exit the wizard. A progress bar will appear while the reports are generated, and you will be prompted to enter a file name and location for the cluster report, export file, and recommendation summary.



Viewing Results with Reports and Tables

Cluster Report

- In this example, 19 tabs are created in the report. The first tab is a summary of all the run and variable information. General information about the run is listed on the first tab, such as the number of variables and respondents and whether the data was standardized or centered.
- Two tabs are created for each cluster size. An 'info' tab and a detailed analysis tab with information about the selected replication. All the replications and specifications used in the run are listed for reference at the top of the report.

Number	Tab Label Name
1	Summary
2	Cluster 2 Info
4	Cluster 3 Info
5	Cluster 3 Repl 5
6	Cluster 4 Info
7	Cluster 4 Repl 8
8	Cluster 5 Info
9	Cluster 5 Repl 4
10	Cluster 6 Info
11	Cluster 6 Repl 4
12	Cluster 7 Info
13	Cluster 7 Repl 6
14	Cluster 8 Info
15	Cluster 8 Repl 7
16	Cluster 9 Info
17	Cluster 9 Repl 9
18	Cluster 10 Info
19	Cluster 10 Repl 4

Imported Data				
	Percent	Included	Excluded	Total
Respondents	95.64%	526	24	550
Attributes	80.00%	28	7	35

Clustering Methods			
Full Method Name	Replication	Clusters	Iterations
IK Means-simple (Forgy) Distance based	1	2	14
IK Means-simple (Forgy) Distance based	1	3	15
IK Means-simple (Forgy) Distance based	1	4	15
IK Means-simple (Forgy) Distance based	1	5	14
IK Means-simple (Forgy) Distance based	1	6	13
IK Means-simple (Forgy) Distance based	1	7	21
IK Means-simple (Forgy) Distance based	1	8	27
IK Means-simple (Forgy) Distance based	1	9	45

Variable	Cluster 2 Info	Cluster 2 Repl 4	Cluster 3 Info	Cluster 3 Repl 6	Cluster 4 Info
14.03	-3.46	-3.11	5.44	11.80	
9.62	-2.97	0.20	2.61	4.35	
9.72	-3.33	-0.55	3.48	5.69	
17.48	-3.02	4.75	-0.59	4.76	
17.88	-4.49	-3.17	6.44	12.74	
7.60	-2.02	-0.93	2.55	3.49	
8.90	-2.74	-2.20	4.12	6.55	
41.13	-24.30	19.67	8.45	153.34	
62.41	-4.89	0.06	4.49	7.26	
37.72	-2.50	7.03	-2.72	6.96	
29.00	4.18	-3.21	-1.57	3.34	
37.86	-15.07	7.66	8.50	40.11	

Viewing Results with Reports and Tables

Cluster Deviations

- The example table on the right is the seventh tab in the report file and shows the mean deviations for the four cluster solution.
- Scroll down to the deviation table, the second table in the tab. This is the easiest way to identify differences between the clusters.
- The name and size (number of cases) of each cluster are listed across the top of the table.
- Note, in the example at right, problem resolution is most important to the third cluster and online service is very important to the second cluster. For more detailed information about the report or selecting other options for the cluster analysis, please refer to the SegmentSolve™ manual.

Report - MA_DemoReport.xls

Means as Deviations from Grand Means and F Ratios									
	Cluster Number	Grand Mea	1	2	3	4	F Ratio		
	Cluster Name		CI1	CI2	CI3	CI4			
	Cluster Size								
44	1	IMPRSPD	Problem Resolution Speed	40.04	-9.05	-11.89	15.88	0.96	33.27
45	2	IMPRCTCT	Ease of Problem Resolution	37.16	-12.63	-9.73	16.74	3.25	48.16
46	3	IMPRINT	Interest in Solving Problems	39.82	-11.50	-2.15	13.44	2.27	35.32
47	4	IMPRFU	Problem Resolution Follow Up	27.67	-11.43	-7.97	20.92	-1.52	52.05
48	5	IMPRDOFF	Breadth of Offering	12.64	-1.01	-8.14	11.19	-4.73	19.22
49	6	IMSUPPT	Staff Support Type	19.03	-3.12	-4.07	15.29	-6.85	35.13
50	7	IMPROACT	Banker Contact Frequency	19.65	1.10	-7.32	11.92	-7.61	21.61
51	8	IMUNDR	Staff Understands Needs	34.43	-13.56	-11.16	14.95	5.90	54.16
52	9	IMKNOW	Staff Provide Correct Info/Ans	42.28	-13.31	-0.77	11.42	5.12	40.26
53	10	IMOPNFU	Follow up After Account Opening	20.69	-2.13	-1.82	11.08	-5.43	18.34
54	11	IMQCKTX	Quick and Efficient Transactions	38.40	-13.93	0.25	7.05	8.63	38.47
55	12	IMAINSPD	Account Information Speed	25.70	-5.12	-6.66	11.91	-1.61	17.59
56	13	IMSTMNT	Account Statement Type	18.00	-1.80	-2.82	10.38	1.94	17.71
57	14	IMERRORS	Transaction Errors in Past Year	53.05	-13.40	-4.31	-3.13	16.90	41.55
58	15	IMPROFNL	Professional Manner	36.63	-11.41	-6.20	7.99	7.36	25.76
59	16	IMHOURS	Branch Hours	27.28	3.76	-9.62	9.86	-8.06	18.06
60	17	IMBOCBOC	Min. Balances Pref. Rates, etc.	19.27	-0.61	3.42	3.61	-3.05	2.40
61	18	IMBRWTTM	Branch Wait Time	21.42	-1.08	-11.45	8.40	-1.64	10.03
62	19	IMATMWT	ATM Wait Time	14.28	6.60	4.30	-1.32	-6.89	13.54
63	20	IMHLDTM	Hold Time on Phone	9.68	3.75	5.12	-0.27	-1.70	4.46
64	21	IMVEBSPD	Online Response Speed	9.87	-6.03	41.34	-2.29	-4.69	139.23
65	22	IMTRVBR	Travel Time to Branch	17.56	1.54	-0.16	5.56	-5.54	6.12
66	23	IMTRVATM	Travel Time to ATM	17.75	5.77	8.62	-1.75	-7.04	11.64
67	24	IMPHNEED	Phone Banking Meets Needs	7.53	3.17	-2.50	-1.18	-1.55	2.71
68	25	IMVEBND	Online Banking Meets Needs	8.91	-6.15	52.66	-4.93	-6.02	398.40
69	26	IMFEEC	Monthly Fee-Checking	61.81	5.32	2.01	0.40	-6.18	6.49
70	27	IMRATEC	Interest Rate-Checking	29.41	4.10	-8.82	2.21	-3.06	3.76
71	28	IMATFEE	Fee for Using Foreign ATM	37.63	9.73	1.88	5.78	-14.45	22.45

Viewing Results with Reports and Tables

■ Recommendation Summary

- The recommendations report is the main decision support report produced by SegmentSolve™. It shows the best cluster solution and then characterizes those segments using the categorical variables that were selected in Step Five of the Wizard. Thus it not only tells the analyst “What the Segments Are” (the middle part of the report), but it also tells the analyst “Who/How” those segments are characterized (the bottom part of the report.)
- There are three sections to the report.
 - Header:** The name and description of the cluster solution and the directory location of the configuration file (.SSC file) is listed.
 - What the Segments Are:** The cluster variables are shown as deviations from the grand mean for each cluster. The cluster variables are in file order, and the deviations that are significantly different from the grand mean are highlighted.
 - Who They are:** All of the subgroup variables that were selected are cross-tabbed by the cluster variables. Next segment levels that are significant present or significantly absent from that cluster (as measured by the standardized residual statistic) are shown. (A number greater than 1.7 means that that group is found significantly present in that cluster. A number less than -1.7 means that that group is significantly absent from that cluster.)

The screenshot shows an Excel spreadsheet titled 'MA_TrialRecommend.xls'. The report is organized into several sections:

- Segmentation Summary:** Located at the top, it includes fields for Name (Solution 46 [6,2]), File (C:\Program Files\Market Advantage\SegmentSolve\MA_Trial.SAV), Date (Tue Oct 15, 2002), and Time (13:58:58). A summary text states: "The segmentation being summarized is the two group solution using the TK Means-reflected (Jancey) method, using Centered data."
- Cluster Headers:** Two columns represent Cluster C1 and Cluster C2. Each header includes a cluster name and size (C1: 265.00, C2: 200.00).
- What the Segments Are:** This section contains two tables. Each table lists variables and their standardized deviations from the grand mean. Significant deviations are highlighted in purple.

Variable Label	GrdDev	Variable Label	GrdDev
Problem Resolution Speed	0.627	Problem Resolution Speed	-0.831
Ease of Problem Resolution	1.817	Ease of Problem Resolution	-2.408
Interest in Solving Problems	1.957	Interest in Solving Problems	-2.593
Problem Resolution Follow Up	-2.034	Problem Resolution Follow Up	2.695
Product Range	-4.582	Product Range	6.071
Staff Support Type	-5.831	Staff Support Type	7.726
Banker Contact Frequency	-8.740	Banker Contact Frequency	11.581
Staff Friendliness	0.150	Staff Friendliness	-0.199
Staff Knowledge	4.160	Staff Knowledge	-5.512
Follow up After Account Opening	-3.890	Follow up After Account Opening	5.155
Quick and Efficient Transactions	3.721	Quick and Efficient Transactions	-4.930
Account Information Speed	-2.599	Account Information Speed	3.443
Account Statement Type	-5.720	Account Statement Type	7.579
Transaction Errors in Past Year	16.663	Transaction Errors in Past Year	-22.078
Professional Manner	4.301	Professional Manner	-5.698
- Who They Are:** This section contains two tables showing standardized residuals for various variables.

Variable/Level Label	Std.Res.	Variable/Level Label	Std.Res.
Overall Satisfaction with Provider		Overall Satisfaction with Provider	
Slightly Satisfied	-1.742	Slightly Satisfied	2.005
BALANCE		BALANCE	
No Balance	-15.849	No Balance	-13.718

For more information contact:

Decision Support Sciences, Inc., East West Technology Center
1751 W. Diehl Road, Suite 150, Naperville, IL 60563
www.DecisionSupportSciences.com
(630) 428-1847

To download the complete 150 page manual for SegmentSolve™, go to
the following link:

www.DecisionSupportSciences.com/downloads.html