Interactive Analysis of Survey Data Using SAS/AF and SAS/EIS Software


ABSTRACT
The National Agricultural Statistics Service (NASS) conducts numerous surveys to gather and disseminate agricultural statistics. NASS's Agricultural Statistics Board (ASB) is responsible for setting estimates for a variety of agricultural commodities. Currently the process of reviewing survey data and other information in setting estimates is a heavily paper oriented process. Shuffling through papers to review and analyze data can be a hurried and frustrating process when the time to publish estimates is limited.

With SAS/EIS and SAS/AF Frame software you can build practical information systems using object oriented programming. These applications are graphically oriented and allow you to navigate through the application, making choices and analyzing data simply by pointing and clicking with a mouse.

A prototype executive information system was created to incorporate some of the procedures and information used by the ASB in setting estimates. This prototype was developed with two objectives in mind: (1) for the author to become familiar with the SAS/EIS and SAS/AF software, and (2) to determine if NASS could use the software to develop practical applications to replace or augment specific procedures currently conducted in the review and analyses of survey data.

1. INTRODUCTION
The National Agricultural Statistics Service (NASS), an agency of the U.S. Department of Agriculture, collects, compiles, and disseminates current national and state agricultural statistics. Its key mission is to provide accurate and reliable agricultural forecasts and estimates. Through surveys, field observations, objective yield measurements, and other administrative data, NASS's Agricultural Statistics Board (ASB) annually publishes estimates for about 120 crops and 45 livestock items in approximately 350 reports. Some of these commodities include inventories and production of hogs, cattle, sheep and wool, goats, catfish, trout, poultry, eggs, and dairy products; acreage, yield, and production of grains, hay, oilseeds, cotton, tobacco, some fruits and vegetables, floriculture, and other specialty crops; and economic data related to agriculture.

In setting estimates at the national level, the members of the ASB consider a huge quantity of information. While some of the data is contained in a computer spreadsheet, a large portion of the data, such as charts, tables, and data listings, is also distributed to each ASB member in hardcopy form. The resources required to develop and distribute multiple copies of such materials can be considerable. The review of all this information is a tedious and time consuming process. With an executive information system as described in this paper, all information would be accessible at the click of a mouse. The flexibility of maintaining pieces of the system could save considerable time and resources. For example, consider the short time required to change the appearance of a graph by modifying a few lines of SCL code of an AF FRAME entry compared to the time to recreate a graph and to print multiple paper copies.

In conducting its surveys, NASS uses current statistical methodology to obtain the data and produce estimates within days of the completion of each survey. To ensure that every farm and every acre of land have a known probability of being selected into the sample, NASS uses a multiple frame approach to its sample design. One frame involves a sample of land area segments and the other uses samples from a list of farms and ranches stratified by size. For hog inventory, the portion of the estimate that is due to the area frame (also referred to as NOL in this paper) is of particular interest. In recent surveys, it has been noticed that a small number of records from the area frame can have a large influence on the precision of the estimate. This is in part due to the fact that, in general, samples from the area frame have larger survey weights than the list samples. These large weights can result in outliers. While outliers can also occur for list records, list outliers are generally due to reported data and not survey weight. One feature of this prototype system is the ability to identify potential outliers and to see how these observations affect the estimates of hog inventory.

The purpose of the system is to show how the ASB data review process could be integrated into one user friendly application that allows users to interactively analyze survey data for the current quarter in a variety of formats. This system is not meant to replace ASB procedures, only to show what could be done using SAS/EIS and SAS/AF software. This system uses fictitious data to demonstrate some of the analyses that might be considered when setting estimates. Summarized data for different estimators are compared over time in line graphs and bar charts; estimates can be entered by editing a SAS data set; the effect of outliers can be analyzed through the use of maps, graphs, and tables; ad-hoc in-depth analysis can be performed by invoking SAS/INSIGHT; notes pertaining to past surveys can be viewed; and summarized data for different estimators can be compared in spreadsheets using SAS/CALC.

The system described in this paper uses SAS release 6.08 products SAS/AF, SAS/CALC, SAS/INSIGHT, and Screen Control Language to provide user-friendly point-and-click applications for accessing and analyzing survey data in a variety of formats including graphs, spreadsheets, and tables.
II. SYSTEM DESCRIPTION

A. Main Menu

The Main Menu (see Figure 1) was created using the EIS Desktop object, which lets the developer build menus using graphics to represent choices to end-users. The first step in creating the Main Menu was to use SAS/GRAPI software to create three U.S. maps, each used to depict a different level of access to the data. These graphs were stored in a catalog and then used in the menu via the SAS/GRAPH Output widget, which displays the contents of a GRSEG catalog entry in the region occupied by the widget. Next, the Target Application attribute was defined for each map. Target applications specified in the Target Application window execute when the user clicks on the region being defined. Each map in the Main Menu has as a target an EIS Block Menu, representing choices for the specific level that was selected.

![Main Menu](image1)

A SAS software supplied icon was used for the EXIT icon. The target application for this icon is an Exit Menu, which terminates the application when the icon is selected. The Text Label widget was used to provide the labels under the maps and the EXIT icon, and the instructions at the bottom of the screen. Graphic Text was used to create the menu title.

The default EIS commands HOTSPOT, NOTES, BOOKMARK, GOBACK, and HELP, usually represented by push buttons at the bottom of the screen, were not all applicable to this menu. A customized command menu was generated using PROC PMENU to include only the options GOBACK and HELP with a customized HELP catalog entry. The name of this menu was specified in the General Attributes window of the Desktop object.

![Block Menu](image2)

Data for the present survey can be accessed at three different levels, as indicated by the maps. The options provided by the target EIS Block Menus for the three levels are similar. However, due to length constraints of the paper and the presentation, only the U.S. Level options are discussed here.

B. U.S. Level Block Menu

When the user clicks on the "U.S. Level" map in the Main Menu, the EIS Block Menu in Figure 2 is displayed. Block menus provide a simple, attractive means of presenting choices to users. There are five blocks that will allow you to: view the data in a variety of formats including charts, maps, tables, and spreadsheets; edit a SAS data set and set final hog estimates; and view notes stored in a catalog entry. The push button labeled "GOBACK" returns the user to the Main Menu when clicked on.

To define the blocks, the "Edit Menu Text" button is selected in the Block Menu window. This allows you to provide a title for the block and to specify the target application you want to assign to the block. Four of the blocks have an EIS LISTMENU application as the target. A LISTMENU application allows the user to choose from an unlimited number of options from either a "popmenu" or a scrollable list. Figure 3 shows the contents of the "listmenus" for these four blocks.

![Block Menu With Options](image3)
1. CHARTS

Selecting the CHARTS block allows you to dynamically create a wide variety of graphs for viewing the data. When you click on "Inventory Item", the AF Frame entry shown in Figure 4 is displayed.

Figure 4. Inventory Items Default Chart

This frame uses push buttons, "popmenus", and graphics to present options to the user. Extensive use of SCL lists was employed to process user input in an SCL program to update the graph. As you make selections, the line chart, the titles, and axes labels are dynamically updated via the SCL program. This frame is used to compare three estimates for several items over time. The default legend provided with the Graphics object took up too much space. Text labels and a container box were used to make the legend in the lower left hand corner of the frame. The default chart compares the estimates for the total number of hogs in the United States by quarter from March 1989 to September 1993. (Note: These numbers are not the true U.S. level indications set by NASS.)

By clicking on the push button and selecting from the resulting "popmenu", you are able to create graphs for four different items. Also, by selecting a specific quarter from the QTR list, the user can create charts that compare the estimates over five years for a specific quarter. The "GOBACK" push button issues the END command, returning the user to the block menu.

Figure 5 shows the list of choices for each push button. Another feature of this frame is the DATA INFO box in the upper right corner of the frame. What is shown is not the actual default DATA INFO box created by the "SHOW_CLICK_INFO" method. I did not like the appearance of the default DATA INFO box, so an AF PROGRAM entry was created to perform the same function. When you click on a point of the graph, an SCL list is created containing information about that point. This information is then passed to the PROGRAM entry, which displays the values of the point clicked on.

Figure 5. Plot Showing Pop-up Menus

When you select CHARTS => Survey Comparisons, Figure 6 is displayed. This is another FRAME entry consisting of push buttons, "popmenus", and graphics, that uses an SCL program to process user input to update a bar chart. The purpose of this frame is to show the contribution of the area frame samples (NOL) to the survey indications for total hog inventory and pig crop over time. You can view the contribution of the area frame to the total expansion and to the standard error by month for two different estimators - a reweighted estimate (MF RWT) and an adjusted estimate (MF ADJ). The reweighted estimate is an estimate that has been adjusted for non-response (based on the number of usable records from the sample). The adjusted estimate is an estimate based on imputed data. NASS's ASB does not just use the indications from the survey to set estimates, but considers other information from a variety of sources. These charts allow you to see the historical comparison of the estimate set by the ASB in relation to different indications calculated from the survey data.

Figure 6. Charts Default Bar Chart
Figure 7 shows the options for each push button, which are displayed by using SCL lists and the popmenu function. Also shown is another DATA INFO box, which passes parameters to a PROGRAM entry. The DATA INFO box is displayed whenever you click on any bar in the chart. In this example, the June 1993 bar was selected. Again, the GOBACK button returns the user to the previous block menu.

2. BALANCE SHEET
Figure 8 displays the FSEDIT application in which members of the ASB could enter their estimates for total U.S. hog inventory and pig crop for the current quarter.

This application uses a "balance sheet" approach by starting with the previous quarter's inventory (ON HAND). Components are added to get a "Total Supply" and other components representing loss of inventory are totaled to get a "Total Disposition". The difference between supply and disposition is then calculated to get the "Indicated" total. The difference between "Indicated" and "Estimated" is then calculated. The target for the difference is zero, but this is not always achieved as several relationships between the components of the balance sheet must be considered when setting a final estimate. Changes made here represent changes to the data set used to generate the graphs under the CHARTS block. Thus, users can immediately see how their estimates compare to previous estimates for hog and pig crop inventories.

3. ANALYSIS
One of the major features of this system is the ability to examine the effect outliers have on the hog inventory and variance estimates. The block item "ANALYSIS" allows the user to see how outliers influence the estimates and to even view information about the outliers at the individual record level.

When you select "ANALYSIS ==> Outlier Contribution to Expansion", Figure 9 is displayed. This is an AF FRAME entry consisting of a SAS/GRAPH Output object which displays graphics output from a GRSEG catalog entry. The FRAME also has a push button which issues the END command when clicked on, allowing you to return to the previous screen. The graph that is replayed is a map of the 16 states that are considered the major hog producing states in the U.S. Different colors are used to show the percent of the total (state) inventory that is due to data classified as "outliers".
One of the limitations of using the SAS/AF FRAME Graphics Class here is that you are not able to create a logarithmic scale for the axes. Because of a few large "outliers", the log of the data had to be used to avoid clustering most of the data points near the origin. This inconvenience could be overcome by using a PROC GPLOT with the appropriate axes statements within a SUBMIT block in the SCL program to produce the graph. However, I opted to transform the data first and then use the Graphics class so that you could continue the analysis to another level by clicking on any point in the graph to get a data listing of that record and all other records of the same type.

Figure 11 shows the extended table listing that is generated when a specific type of record is selected. In this example, a NOL OUTLIER was selected. An extended table can be used to display or edit data. In this case, it is used to display the values for six variables from the data set.

Another limitation of the FRAME entry is that when a point is clicked on in a Graphics object, the information returned in the SCL list does not contain the observation number.
The option "ANALYSIS" in SAS/INSIGHT provides more advanced SAS users an opportunity for ad-hoc analysis of the data. Selecting this option invokes SAS/INSIGHT and loads the data for the current survey. Any analysis at this point is entirely up to the user. SAS/INSIGHT will not be discussed in this paper.

4. BOARD NOTES

Selecting the Block "BOARD NOTES" results in a pop-up menu from which the user can choose to view notes pertaining to a particular survey period. The notes contain such information as changes in survey procedures or questionnaires and other conditions that may have influenced current survey. Any analysis at this point is entirely up to the user. SAS/INSIGHT will not be discussed in this paper.

5. COMPARISON TABLES

When the user clicks on the COMP TABLES (comparison tables) block in the U.S. Block Menu, an EIS Listmenu provides four options to choose from (see Figure 3). Selecting an item from the listmenu brings up a SAS/CALC spreadsheet comparing the three estimates for the selected item over time. The columns labeled "% MFR" and "% MFA" in Figure 14 show the ratio of the ASB estimate (BOARD) to the two survey estimates (reweighted and adjusted) as a percent. This is another view of the data that was used to produce the charts under the CHARTS -> INVENTORY ITEM selection. The BOARD estimate for the current quarter are updated by using the values from the data set used in the FSEDIT application.

CONCLUSION

The system presented in this paper is a prototype that the author created in order to become familiar with SAS/EIS and SAS/AF software. I was impressed with the ease at which some applications were developed and the power of SAS/AF and SCL. However, as a new user, I was often frustrated in trying to find information in the usage and reference manuals, especially for SCL lists. I was also a little disappointed in the ability of some of the SAS/EIS features, such as the GRAPHICS class, as mentioned above. It was exciting to use the OOP and GUI features of this software. Overall, I found that SAS/EIS and SAS/AF Frame are very powerful and flexible tools for developing information systems. The Screen Control Language lists were particularly useful for processing user input. I found that using AF Frame with SCL was much more flexible than using SAS/EIS. I found that many of the SAS/EIS objects had limited features and hope that more attributes and features are added in newer versions.

This exercise presented the opportunity to evaluate the possibility of using the SAS System to develop other information systems that would enable users to use the limited time during the survey period to more efficiently analyze survey data to improve survey data quality.

I plan to continue exploring the features of SAS/EIS and SAS/AF software in a project that will build upon the prototype presented here. In fact, a new application has been developed that will be used in the interactive analysis of survey data during the data collection period of one of NASS’s major surveys. This new system will be field tested in March 1995, with hopes of implementing it into the operational program by September 1995. I think that systems developed with SAS/EIS and SAS/AF software can help to more efficiently analyze survey data and to improve survey data quality.

ACKNOWLEDGEMENTS

SAS, SAS/AF, SAS/CALC, SAS/EIS, SAS/FSP, SAS/GRAPH, and SAS/INSIGHT are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. * indicates USA registration.

REFERENCES


