Making the VISUAL Leap From PROC TABULATE to SAS/GRAPH® Software

Dain Crawford

Abstract

PROC GCHART along with the annotate facility allows you to quickly create large quantities of visual reports that can be personalized to accommodate any specifications. When we decided that the output obtained using PROC TABULATE no longer fulfilled our needs, SAS/GRAPH was the logical choice. We have always produced large numbers of personalized reports (as many as 5000 unique pages of reports for one project), but the prospect of translating them into graphs seemed nearly impossible. PROC GCHART, however, actually allowed us to decrease the number of pages we produced while at the same time greatly improving the quality and usability of our product. We are able to control every aspect of the graph for maximum use of the page space.

Introduction

No one can dispute the power of PROC TABULATE for flexibility and capability. Its ability to distill large amounts of data into neatly formatted reports is unquestioned. For internal use, we will continue to use PROC TABULATE when dealing with large amounts of data.

But what about our customers? Do they want to see PROC TABULATE reports of their data? Or do they want to be able to look at results and immediately be able to identify what kind of action they need to take? How do we present their data to them in a way that is clear, concise and speaks for itself? We made the transition from PROC TABULATE to SAS/GRAPH with impressive results.

We process Employee Opinion Surveys for hospitals around the country. Our surveys are usually comprised of about 80 statements that are summarized into ten categories. The data is presented for the hospital overall, for each department within the hospital, as well as for each individual manager. When it is available, we will compare up to three years of data for these areas. On department and manager output, we also include overall hospital data for greater ease in comparing the data. It is not uncommon to generate 3000-5000 pages of reports by the time we have accounted for all of the various breakdowns.

Exhibit 1 is a typical PROC TABULATE report. It contains a great deal of information and is neatly organized. But how readily can the comparisons between statements, years or personal and hospital data be made?

Exhibit 2 is an example of a horizontal bar chart using PROC GCHART. This type of presentation makes it very easy for anyone to compare the data by category and statement from this year to last year. Because the hospital overall data is represented by “dots,” it is also possible to make this further comparison by simply glancing at the page. There is no need to make mental calculations to understand the data.

Input Data

For both the PROC TABULATE and PROC GCHART methods of processing, the data is collected from scanable survey instruments and massaged with SAS to create an input file.

Exhibit 3 is an example of the type of input file that would be set up for using PROC TABULATE. While PROC TABULATE handles the question variable very well and wraps text within a column, to achieve the desired look using SAS/GRAPH, it is necessary to use the annotate facility to place the statements on the page.

Generally, multiple years of data are compared for each survey statement. In order to accomplish this using the PROC GCHART method, the input file used must contain a record for each statement for each year. If we have 10 statements and three years of data for each, the data set going into PROC GCHART must have 30 records as shown in Exhibit 4.

Annotate

While the annotate facility has a tremendous amount of flexibility, it can be difficult to master. The old "flexible therefore complex" syndrome certainly applies here. Once the data is organized, the annotate dataset is created. Annotate allows you to place any text, symbol or graphic element on the page in any configuration you require. This is the very feature that makes using SAS/GRAPH in conjunction with the annotate facility worth the effort. Our survey statements can be up to 190 characters long. Since we prefer to show the entire statement and not just an abbreviation when we feedback data to our customers, we had to find a way to create wrapped text. A data step is used to create the annotate data set so all of the SAS programming tools are available for use. We opted to write a quick parsing routine to break the statement into three lines.

\[ \text{lgh} = \text{length (question)}; \]

\[ \text{do i = 1 to lgh;} \]

\[ \text{qst(i) = substr(question, i, 1);} \]

\[ \text{end;} \]

\[ \text{qa=question;} \]

\[ \text{if lgh > 55 then do;} \]

\[ \text{do i = 55 to 1 by -1 until (flg);} \]

\[ \text{qa=substr(question, i, 1);} \]

\[ \text{qb=substr(question, i, 1);} \]

\[ \text{qs=substr(question, i, 1);} \]

\[ \text{if qst(i) = ' ' then flag=1;} \]

\[ \text{end;} \]

\[ \text{end;} \]

This code parses the entire statement by character and then, starting at position 55, looks backward to find a space. When it finds a space it does a substring from position one to the position i.
the space. That data is line one. Another substring sets up the rest of the data as input to a similar routine to build lines two and three. It's not fancy, but we have found it to be very effective.

Output

The PROC TABULATE code is fairly simple using a "by" variable to generate data for each department or manager.

    proc tabulate order=formatted;
    class category question;
    by codename;
    var meanh3 meanh1 meanh2 m3 m1 m2 diff;

Conclusion

Using PROC GCHART has all number of pages we produce charts in ways we had not previously chart and quickly understand this customers.

Obviously, your decision to choose is upon your individual needs. If you are with little or no personalization, this is more than you need. If, however, consistently producing large quantities of SAS/GRAPH with the annotated...
This report compares years I, II and III managers by category and question for the manager below.

Manager = 009 Sue

<table>
<thead>
<tr>
<th>Category</th>
<th>HOSP</th>
<th>HOSP</th>
<th>HOSP</th>
<th>PERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperation</td>
<td>MEAN</td>
<td>MEAN</td>
<td>MEAN</td>
<td>MEAN</td>
</tr>
</tbody>
</table>
This Graph Compares Year I Means, Year II Means and Year III Means by Statement
The means range from 5.00 (highest) to 1.00 (lowest)

Manager = 009 Sue

Cooperation
<table>
<thead>
<tr>
<th>CODE</th>
<th>CODENAME</th>
<th>CATEGORY</th>
<th>QUEST</th>
<th>MEANH3</th>
<th>MEANH2</th>
<th>MEANH1</th>
<th>M3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>001</td>
<td>John</td>
<td>1</td>
<td>10</td>
<td>3.78</td>
<td>3.68</td>
<td>3.59</td>
</tr>
<tr>
<td>2</td>
<td>001</td>
<td>John</td>
<td>1</td>
<td>13</td>
<td>4.03</td>
<td>4.02</td>
<td>4.01</td>
</tr>
<tr>
<td>3</td>
<td>001</td>
<td>John</td>
<td>1</td>
<td>14</td>
<td>3.57</td>
<td>3.59</td>
<td>3.62</td>
</tr>
<tr>
<td>4</td>
<td>001</td>
<td>John</td>
<td>1</td>
<td>16</td>
<td>3.87</td>
<td>3.82</td>
<td>3.70</td>
</tr>
<tr>
<td>5</td>
<td>001</td>
<td>John</td>
<td>1</td>
<td>24</td>
<td>3.25</td>
<td>3.14</td>
<td>3.01</td>
</tr>
<tr>
<td>6</td>
<td>001</td>
<td>John</td>
<td>1</td>
<td>51</td>
<td>4.20</td>
<td>4.20</td>
<td>4.14</td>
</tr>
<tr>
<td>7</td>
<td>001</td>
<td>John</td>
<td>1</td>
<td>52</td>
<td>4.09</td>
<td>3.96</td>
<td>3.86</td>
</tr>
<tr>
<td>8</td>
<td>001</td>
<td>John</td>
<td>1</td>
<td>59</td>
<td>3.53</td>
<td>3.49</td>
<td>3.33</td>
</tr>
<tr>
<td>9</td>
<td>001</td>
<td>John</td>
<td>1</td>
<td>88</td>
<td>3.79</td>
<td>3.74</td>
<td>3.66</td>
</tr>
</tbody>
</table>