Paper 369-2013

Using Design Principles to Make ODS Template Decisions

Helen Smith, RTI International, Research Triangle Park, NC;
Susan Myers, RTI International, Research Triangle Park, NC

ABSTRACT

With the Output Delivery System (ODS), SAS® continues to provide programmers with many style templates for developing reports. These default templates and style definitions present the data in a clear and attractive manner often with no further thought needed. The website "SAS ODS Style Examples" provides a link to screen shots of SAS ODS style templates (Swoope, 2011). However, when producing complicated reports with multiple requirements, using basic design principles to determine which template or which custom style definition to use can make for a more readable and comprehensive final report.

This paper presents the code and the design considerations for two ODS reports; one, a redesign of a 10-plus year old SAS® program originally designed with PUT statements, and two, a highly customized SAS® program for delivering output in Excel.

The reports are generated using Proc Report and developed with version 9.2 and 9.3 of SAS® executing on a Windows XP platform. The reports use features that are not platform specific.

INTRODUCTION

Perceptually, we notice the differences and similarities of what we see in order to navigate and make sense of the world. How we do this was first examined in detail in the 1920s and 1930s by a group of researchers known as Gestalt psychologists, "gestalt" being the German word for "shape" or "form."

We differentiate between what is figure/object (important) and what is not background/ground (not so important) and group what we see into a hierarchy of visual and conceptual importance. Those elements that we see as figure (important) tend to be dominant, darker, larger, more colorful, and so on. Everything else appears less dominant and less important.

For those elements (numbers, shapes, lines, color, or the data that make up a SAS® report) that appear to be similar, we unify visually and think of them as a group (principle of similarity). For an element that looks different from a larger group of similar shapes will seem separate from that larger group. For elements that are close to each other, we will associate them with one another even if the elements are quite different (the principle of proximity). The farther apart these elements are, the more tenuous the visual association.

In Figure 1, the first figure on the left shows three white circles and a white rectangle in a dark, gray rectangle. We group the three circles together because they are similar and close together. We see the rectangle apart from the group. The white rectangle is different, yet at the same time we notice the overarching form of a circle created by the white circles and rectangle (principles of continuity and closure).
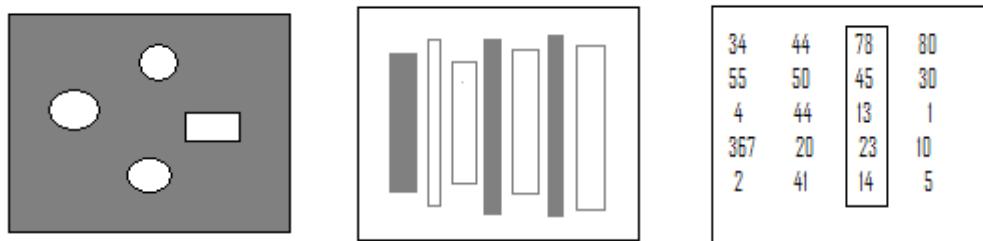

Figure 1. Similarity, Proximity, Continuity, Closure

In the second figure, we see seven vertical stripes of varying thicknesses. Three are dark and form a visually dominant group. The other white verticals form another less visually dominant group: they are unified by the similar vertical orientation, proportion, shape, and relative closeness or proximity of the verticals.

Finally, on the far right, we see four columns of numbers. The numbers are close to one another, are similar, and form a recognizable form (a column) bracketed by white space on either side. One column looks different from the other columns. We notice it before we see the other columns of numbers because it is outlined. By outlining it—and thus making it different—we have designed that column to be seen as important and visually dominant.

In the following paragraphs, we discuss how these concepts were applied to the redesign of two SAS® programs. The first redesign is *additive,* slowly adding features to complete the report. The second redesign is *subtractive,* taking away original features to redesign the report.

REDESIGN OF A LEGACY SUMMARY STATUS REPORT

Our first report is a legacy program written in SAS® using arrays and PUT statements. It produces a listing of all the production status codes for a telephone survey, including the call counts for the current and previous weeks. Its purpose is to provide project management with a daily and weekly snapshot of production call activity.

Figure 2 shows output from the report and the PUT statements used to generate the output for the first two tables.

```
Project X Summary Status Report for Wednesday, May 23, 2012
                    Daily and Weekly Summaries

                                 Daily     Weekly    Week to    Cumulative
                                           (7 days)  date
    ------------------------------------------------------------------------
    Interview pending               0         4         0           28
    Final ineligible                0         0         0            5
    Final unable to locate          0         0         0          153
    Completed interview             0      1943         0        49319
    Hard Copy received              0         0         0            0
    Partial interview               0         0         0            0
    Other final noninterview        0         0         0            0
    ------------------------------------------------------------------------
    Total                                                        49505
    ------------------------------------------------------------------------

    Response rates

         Comp/all final elig                   99.69
         Comp/final elig+pending               99.63
    ========================================================================


                         Status Report

    ========================================================================
    Pending                              28
    ========================================================================
    Active                                        28
    ------------------------------------------------------------------------
    101  Not Ever Released                        0
    102  Released, no action                     20
    105  Obtained New Info                        0
    110  Attempting Preloaded Info                0
    111  Attempting Added Info                    0
    115  Attempting Da/FD Info                    0
    116  Attempting Other New Info                0
    210  No contact, no info left, few attempts   0
    211  No contact, no info left, many attempts  0
    283  Final Refusal --hostile                  0
    285  Language barrier--Spanish                0
    286  Language barrier--other                  0
    ------------------------------------------------------------------------
         Sub-total                               0
    ========================================================================
    Complete or Partial Interviews      49319
    ------------------------------------------------------------------------
    290  Partial interview                        0
    291  Phone Not Answered                      26
    292  Hold Q - 10 Minute Wait                169
    293  Call Dropped                           859
    294  Refused                                  0
    295  COMPLETE - REACHED CSR               46345
    296  Unexpected Breakoff                      0
    297  Disconnect after 2/3 rings             177
    298  IVR - 10 Minute Wait                    92
    299  TI Left IVR                           1651
    ------------------------------------------------------------------------
         Sub-total                            49319
    ========================================================================
    sum_stat MISSING                     0
    ========================================================================
    TOTAL                            49505
    ========================================================================
```

Figure 2. Legacy Summary Status Report

A recent project required that this report display data for the entire sample plus seven subsamples. Tweaking the PUT statement and creating seven more vertical columns might have been sufficient, but using the features of ODS instead produced a better report.

The report has six sections: a table of the overall call counts based on the activity for the current and previous weeks; two tables, with one for each response rate; a table of counts of pending status codes with subtotals; a table of counts of final status codes with subtotals and totals; and an error table.

The initial task was to split the data sets into six separate tables for each of the report's main sections. The ODS style chosen for the report was "normal":

```
ODS listing close;
ODS HTML BODY=stat_rpt Style=normal;
```

This default style is simple, and the output easily fits on one to two 8.5" x 11" sheets of paper if needed.

DESIGNING THE DAILY AND WEEKLIES TABLE

Figure 3 shows the first report table before adding additional bolding or shading/color. Although the header data are bold and in color, the report cells with values and the rows of descriptions disappear in the "forest" of columns and rows.

	Daily	1	2	3	4	5	6	7		Weekly	1	2	3	4	5	6	7		Week to Date	1	2	3	4	5	6	7		Cumulative	1	2	3	4	5	6	7
Interview Pending	75	3	1	0	71	0	0	0	Interview Pending	276	5	10	0	261	0	0	0	Interview Pending	75	3	1	0	71	0	0	0	Interview Pending	614	90	102	154	268	0	0	0
Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0
Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0
Completed Interview	18	0	2	0	16	0	0	0	Completed Interview	46	1	16	0	29	0	0	0	Completed Interview	18	0	2	0	16	0	0	0	Completed Interview	265	154	82	0	29	0	0	0
Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0
Other Final Interview	0	0	0	0	0	0	0	0	Other Final Interview	1	0	0	0	1	0	0	0	Other Final Interview	0	0	0	0	0	0	0	0	Other Final Interview	3	2	0	0	1	0	0	0
TOTAL	93	3	3	0	87	0	0	0	TOTAL	323	6	26	0	291	0	0	0	TOTAL	93	3	3	0	87	0	0	0	TOTAL	882	246	184	154	298	0	0	0

Figure 3. Report Table With No Bolding Except for the Header Text

Figure 4 shows the report after the rows of descriptions are bolded. When bolded, the descriptions take precedence over the columns of data that are lighter and become not only more dominant but also a stronger organizational element in the report.

	Daily	1	2	3	4	5	6	7		Weekly	1	2	3	4	5	6	7		Week to Date	1	2	3	4	5	6	7		Cumulative	1	2	3	4	5	6	7
Interview Pending	0	0	0	0	0	0	0	0	**Interview Pending**	210	2	11	0	197	0	0	0	**Interview Pending**	210	2	11	0	197	0	0	0	**Interview Pending**	632	90	104	154	284	0	0	0
Final Ineligible	0	0	0	0	0	0	0	0	**Final Ineligible**	0	0	0	0	0	0	0	0	**Final Ineligible**	0	0	0	0	0	0	0	0	**Final Ineligible**	0	0	0	0	0	0	0	0
Final Unable To Locate	0	0	0	0	0	0	0	0	**Final Unable To Locate**	0	0	0	0	0	0	0	0	**Final Unable To Locate**	0	0	0	0	0	0	0	0	**Final Unable To Locate**	0	0	0	0	0	0	0	0
Completed Interview	0	0	0	0	0	0	0	0	**Completed Interview**	30	1	16	0	13	0	0	0	**Completed Interview**	30	1	16	0	13	0	0	0	**Completed Interview**	247	154	80	0	13	0	0	0
Partial Interview	0	0	0	0	0	0	0	0	**Partial Interview**	0	0	0	0	0	0	0	0	**Partial Interview**	0	0	0	0	0	0	0	0	**Partial Interview**	0	0	0	0	0	0	0	0
Other Final Interview	0	0	0	0	0	0	0	0	**Other Final Interview**	1	0	0	0	1	0	0	0	**Other Final Interview**	1	0	0	0	1	0	0	0	**Other Final Interview**	3	2	0	0	1	0	0	0
TOTAL	0	0	0	0	0	0	0	0	**TOTAL**	241	3	27	0	211	0	0	0	**TOTAL**	241	3	27	0	211	0	0	0	**TOTAL**	882	246	184	154	298	0	0	0

Figure 4. Report Table With Added Bolding for Row Descriptors

Next, the report table has cells with values greater than zero that are bolded for emphasis (Figure 5). As more cells have values greater than zero, this bolding of the cells becomes less useful as the rows of descriptions become visually more similar and tend to group together. Additional tweaking is needed.

	Daily	1	2	3	4	5	6	7		Weekly	1	2	3	4	5	6	7		Week to Date	1	2	3	4	5	6	7		Cumulative	1	2	3	4	5	6	7
Interview Pending	0	0	0	0	0	0	0	0	**Interview Pending**	210	2	11	0	197	0	0	0	**Interview Pending**	210	2	11	0	197	0	0	0	**Interview Pending**	632	90	104	154	284	0	0	0
Final Ineligible	0	0	0	0	0	0	0	0	**Final Ineligible**	0	0	0	0	0	0	0	0	**Final Ineligible**	0	0	0	0	0	0	0	0	**Final Ineligible**	0	0	0	0	0	0	0	0
Final Unable To Locate	0	0	0	0	0	0	0	0	**Final Unable To Locate**	0	0	0	0	0	0	0	0	**Final Unable To Locate**	0	0	0	0	0	0	0	0	**Final Unable To Locate**	0	0	0	0	0	0	0	0
Completed Interview	0	0	0	0	0	0	0	0	**Completed Interview**	30	1	16	0	13	0	0	0	**Completed Interview**	30	1	16	0	13	0	0	0	**Completed Interview**	247	154	80	0	13	0	0	0
Partial Interview	0	0	0	0	0	0	0	0	**Partial Interview**	0	0	0	0	0	0	0	0	**Partial Interview**	0	0	0	0	0	0	0	0	**Partial Interview**	0	0	0	0	0	0	0	0
Other Final Interview	0	0	0	0	0	0	0	0	**Other Final Interview**	1	0	0	0	1	0	0	0	**Other Final Interview**	1	0	0	0	1	0	0	0	**Other Final Interview**	3	2	0	0	1	0	0	0
TOTAL	0	0	0	0	0	0	0	0	**TOTAL**	241	3	27	0	211	0	0	0	**TOTAL**	241	3	27	0	211	0	0	0	**TOTAL**	882	246	184	154	298	0	0	0

Figure 5. Report Table With Bolding for Cells With Values Greater Than Zero

Applying a different shade/color for the columns (1, 2, 3, 4, 5, 6, 7) can make the columns less similar yet provides a reasonable visual difference (Figure 6). The resulting pattern of alternating gray and white stripes, however, undermines the more fundamental conceptual and visual grouping set up previously, the columns of descriptions and the columns containing the data.

Figures 7 and 8 are variations on this theme. However, each introduces "visual noise"/"data imprisonment" (Tufte, 1990, p. 64) rather than clarification. The darker or colorful columns distract because they disrupt the order set up earlier. The bars of columns corral the data and make no sense as the strongest visual elements in the table, more dominant than the data in the cells, the rows of description, and the header. The goal is to create a difference—but only enough of a difference to convey the information.

The colored columns fare no better (Figure 8).

	Daily	1	2	3	4	5	6	7		Weekly	1	2	3	4	5	6	7		Week to Date	1	2	3	4	5	6	7		Cumulative	1	2	3	4	5	6	7
Interview Pending	75	3	1	0	71	0	0	0	Interview Pending	276	5	10	0	261	0	0	0	Interview Pending	75	3	1	0	71	0	0	0	Interview Pending	614	90	102	154	268	0	0	0
Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0
Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0
Completed Interview	18	0	2	0	16	0	0	0	Completed Interview	46	1	16	0	29	0	0	0	Completed Interview	18	0	2	0	16	0	0	0	Completed Interview	265	154	82	0	29	0	0	0
Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0
Other Final Interview	0	0	0	0	0	0	0	0	Other Final Interview	1	0	0	0	1	0	0	0	Other Final Interview	0	0	0	0	0	0	0	0	Other Final Interview	3	2	0	0	1	0	0	0
TOTAL	93	3	3	0	87	0	0	0	TOTAL	323	6	26	0	291	0	0	0	TOTAL	93	3	3	0	87	0	0	0	TOTAL	882	246	184	154	298	0	0	0

Figure 6. Report Table With Alternating Gray Stripes

	Daily	1	2	3	4	5	6	7		Weekly	1	2	3	4	5	6	7		Week to Date	1	2	3	4	5	6	7		Cumulative	1	2	3	4	5	6	7
Interview Pending	75	3	1	0	71	0	0	0	Interview Pending	276	5	10	0	261	0	0	0	Interview Pending	75	3	1	0	71	0	0	0	Interview Pending	614	90	102	154	268	0	0	0
Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0
Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0
Completed Interview	18	0	2	0	16	0	0	0	Completed Interview	46	1	16	0	29	0	0	0	Completed Interview	18	0	2	0	16	0	0	0	Completed Interview	265	154	82	0	29	0	0	0
Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0
Other Final Interview	0	0	0	0	0	0	0	0	Other Final Interview	1	0	0	0	1	0	0	0	Other Final Interview	0	0	0	0	0	0	0	0	Other Final Interview	3	2	0	0	1	0	0	0
TOTAL	93	3	3	0	87	0	0	0	TOTAL	323	6	26	0	291	0	0	0	TOTAL	93	3	3	0	87	0	0	0	TOTAL	882	246	184	154	298	0	0	0

Figure 7. Report Table With Alternating, Darker Gray Stripes

	Daily	1	2	3	4	5	6	7		Weekly	1	2	3	4	5	6	7		Week to Date	1	2	3	4	5	6	7		Cumulative	1	2	3	4	5	6	7
Interview Pending	75	3	1	0	71	0	0	0	Interview Pending	276	5	10	0	261	0	0	0	Interview Pending	75	3	1	0	71	0	0	0	Interview Pending	614	90	102	154	268	0	0	0
Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0
Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0
Completed Interview	18	0	2	0	16	0	0	0	Completed Interview	46	1	16	0	29	0	0	0	Completed Interview	18	0	2	0	16	0	0	0	Completed Interview	265	154	82	0	29	0	0	0
Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0
Other Final Interview	0	0	0	0	0	0	0	0	Other Final Interview	1	0	0	0	1	0	0	0	Other Final Interview	0	0	0	0	0	0	0	0	Other Final Interview	3	2	0	0	1	0	0	0
TOTAL	93	3	3	0	87	0	0	0	TOTAL	323	6	26	0	291	0	0	0	TOTAL	93	3	3	0	87	0	0	0	TOTAL	882	246	184	154	298	0	0	0

Figure 8. Report Table With Alternating Colored Stripes

In Figure 9, the report returns to the lighter gray and groups the columns of data into blocks, resulting in a better solution. There is a clear hierarchy established between the header, the rows of descriptions, the blocks of data, and cells with values greater than zero.

	Daily	1	2	3	4	5	6	7		Weekly	1	2	3	4	5	6	7		Week to Date	1	2	3	4	5	6	7		Cumulative	1	2	3	4	5	6	7
Interview Pending	75	3	1	0	71	0	0	0	Interview Pending	276	5	10	0	261	0	0	0	Interview Pending	75	3	1	0	71	0	0	0	Interview Pending	614	90	102	154	268	0	0	0
Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0
Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0
Completed Interview	18	0	2	0	16	0	0	0	Completed Interview	46	1	16	0	29	0	0	0	Completed Interview	18	0	2	0	16	0	0	0	Completed Interview	265	154	82	0	29	0	0	0
Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0
Other Final Interview	0	0	0	0	0	0	0	0	Other Final Interview	1	0	0	0	1	0	0	0	Other Final Interview	0	0	0	0	0	0	0	0	Other Final Interview	3	2	0	0	1	0	0	0
TOTAL	93	3	3	0	87	0	0	0	TOTAL	323	6	26	0	291	0	0	0	TOTAL	93	3	3	0	87	0	0	0	TOTAL	882	246	184	154	298	0	0	0

Figure 9. Simplify by Reducing the Number of Elements for Better Results

THE COMPLETED REDESIGNED SUMMARY STATUS REPORT

For the other tables in the report, using alternating light-gray columns, bolding for the active cells and subheadings was the solution (Figure 10).

Project X BY SAMPLE Summary Status Report : July 31, 2012 12:14:29

Daily and Weekly Summaries

	Daily	1	2	3	4	5	6	7		Weekly	1	2	3	4	5	6	7		Week to Date	1	2	3	4	5	6	7		Cumulative	1	2	3	4	5	6	7
Interview Pending	75	3	1	0	71	0	0	0	Interview Pending	276	5	10	0	261	0	0	0	Interview Pending	75	3	1	0	71	0	0	0	Interview Pending	614	90	102	154	268	0	0	0
Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0	Final Ineligible	0	0	0	0	0	0	0	0
Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0	Final Unable To Locate	0	0	0	0	0	0	0	0
Completed Interview	18	0	2	0	16	0	0	0	Completed Interview	46	1	16	0	29	0	0	0	Completed Interview	18	0	2	0	16	0	0	0	Completed Interview	265	154	82	0	29	0	0	0
Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0	Partial Interview	0	0	0	0	0	0	0	0
Other Final Interview	0	0	0	0	0	0	0	0	Other Final Interview	1	0	0	0	1	0	0	0	Other Final Interview	0	0	0	0	0	0	0	0	Other Final Interview	3	2	0	0	1	0	0	0
TOTAL	93	3	3	0	87	0	0	0	TOTAL	323	6	26	0	291	0	0	0	TOTAL	93	3	3	0	87	0	0	0	TOTAL	882	246	184	154	298	0	0	0

Comp/All Final Elig RR

Percent	Overall	1	2	3	4	5	6	7
Complete / All final eligible	98.88	98.72	100.00	0.00	96.67	0.00	0.00	0.00

Comp/All Final Elig + Pend RR

Percent	Overall	1	2	3	4	5	6	7
Complete / Final eligible + pending	30.05	62.60	44.57	0.00	9.73	0.00	0.00	0.00

Pending Cases table not displayed due to page limitations. Similar to Final Cases table shown below.

Final Cases

Finalized / Completed	ALL	1	2	3	4	5	6	7
Subject Ineligible								
251 Ineligible	0	0	0	0	0	0	0	0
255 Deceased	0	0	0	0	0	0	0	0
256 Not Ever Released	0	0	0	0	0	0	0	0
258 Invalid Case	0	0	0	0	0	0	0	0
SUBTOTAL	0	0	0	0	0	0	0	0
Subject Not Located								
260 Unable to locate	0	0	0	0	0	0	0	0
262 Time Exhausted	0	0	0	0	0	0	0	0
SUBTOTAL	0	0	0	0	0	0	0	0
Subject Located (Non-Interview)								
270 Unavailable for duration of study	1	0	0	0	1	0	0	0
271 Out of country	1	1	0	0	0	0	0	0
272 Incapable/incapacitated	0	0	0	0	0	0	0	0
273 Institutionalized/incarcerated	0	0	0	0	0	0	0	0
274 Has no Phone	0	0	0	0	0	0	0	0
275 Hearing Impaired	0	0	0	0	0	0	0	0
280 Refusal--unreviewed	1	1	0	0	0	0	0	0
281 Refusal--reviewed	0	0	0	0	0	0	0	0
282 Refusal by other--multiple	0	0	0	0	0	0	0	0
283 Final Refusal --hostile	0	0	0	0	0	0	0	0
285 Language barrier--Spanish	0	0	0	0	0	0	0	0
286 Language barrier--other	0	0	0	0	0	0	0	0
SUBTOTAL	3	2	0	0	1	0	0	0
Completes / Partials								
290 Partial interview	0	0	0	0	0	0	0	0
292 Complete, SPN	0	0	0	0	0	0	0	0
295 Complete	265	154	82	0	29	0	0	0
SUBTOTAL	265	154	82	0	29	0	0	0
Final Cases - Total	ALL	1	2	3	4	5	6	7
TOTAL	268	156	82	0	30	0	0	0

Error

Error Report	ALL	1	2	3	4	5	6	7
ERROR								
999 sum_stat MISSING	0	0	0	0	0	0	0	0

Figure 10. Adding Alternating Gray and White Columns, Bolding for Active Cells and Subheadings for the Rest of the Report

Proc Report Source Code for the Dailies and Weeklies Table

```
ODS listing close ;
ODS HTML BODY=stat_rpt Style=normal ;

PROC REPORT data = overall nowd split='~' style(calldef)=[background=cxFFFFFF];
ods escapechar='^';
title1 "&study Summary Status Report : %sysfunc(date(),worddate.)
%sysfunc(time(),time.)";
title2 " ";
title3 "Daily and Weekly Summaries" ;
title4 "   ";

columns desc daily daily1 daily2 daily3 daily4 daily5 daily6
        daily7
        desc weekly weekly1 weekly2 weekly3 weekly4 weekly5 weekly6
        weekly7
        desc todate todate1 todate2 todate3 todate4 todate5 todate6
        todate7
        desc cum cum1 cum2 cum3 cum4 cum5 cum6 cum7 ;
DEFINE desc / display width=50 format=$50.'           ' style(column) =
{font_weight=bold};
```

6

```
DEFINE daily /  display width=5 format=5. 'Daily' style(column) =
[background=cxFFFFFF];
DEFINE daily1 / display width=5 format=5. '1' style(column)=[background=cxCCCCCC];
  …

DEFINE weekly / display width=5 format=5. 'Weekly' style(column) =
[background=cxFFFFFF];
DEFINE weekly1 / display width=5 format=5. '1' style(column)=[background=cxCCCCCC];
  …

DEFINE todate / display width=5 format=5. 'Week to Date' style(column) =
[background=cxFFFFFF];
DEFINE todate1 / display width=5 format=5. '1' style(column)=[background=cxCCCCCC];
  …

DEFINE cum / display width=5 format=5. 'Cumulative' style(column) =
[background=cxFFFFFF];
DEFINE cum1 / display width=5 format=5. '1' style(column)=[background=cxCCCCCC];
  …

(Code repeated for daily, weekly, todate, and cum.)
```

MS EXCEL WORKSHEET FOR SURVEY PARADATA REPORTING

Our second report is a highly customized MS Excel worksheet designed to present a large amount of paradata in a more readable format.

Survey organizations typically have available much more data on the survey process than can be analyzed in a reasonable amount of time. Some of these "paradata" are potentially very useful for identifying data collection concerns that should be addressed to maximize efficiency and data quality. At the call, sample member, and field interviewer (FI) levels, paradata in the form of a record-of-calls are commonly collected but difficult for supervisors and data quality analysts to utilize in their raw format.

This report was designed to compile raw record-of-calls data into an interactive chart for graphic display. The resulting spreadsheet displays call history data (including the date, time, case ID, interviewer ID, and disposition code) for every call made to a sample member. With the report, a reviewer can trace the pattern of calls made by an interviewer to a single household on a single day and get an overall picture of the progress of the field. In addition to presenting the paradata in a visually appealing manner, specific situations can be highlighted to draw attention to potentially problematic cases.

Figure 11 shows the extensive formatting, including column shading, text colors, and highlighting specific conditions within cells. We have disposition codes with unique colors and show them in the header. Each column represents a portion of the day, daytime, and evening, with shading of evening columns to distinguish from each other. We also have additional special shading for weekends to draw attention to workweek versus weekend periods. The intent of the various column shading is to draw attention to particular times of the day and week that may be more suspicious for reported activity. For example, a large number of completed cases during the weekday morning may be improbable due to respondents being at work or school. In addition to the aforementioned formatting, we also define certain patterns to be particularly worthy of further investigation. Cells that contain predefined patterns are highlighted in yellow.

Before we discuss simplifying the report, one of the particularly interesting challenges to programming the first sample was to color code individual disposition codes within a single cell. The code excerpt provided below (**Source Code to Set Colors**) demonstrates how to assign specific colors to a character variable. The codes that occur within the timeframe represented in a cell are concatenated and separated by commas.

Figure 12 illustrates a cell that contains multiple disposition codes requiring more than one font color. Working in Excel alone, this capability is not available through formatting cells. However, SAS provides the flexibility to accomplish the task by setting each disposition code to a specific color and then concatenating as needed for all codes that occur in a specific time period (i.e., cell).

7

Figure 11. Former Main Interface for SAS Management Console
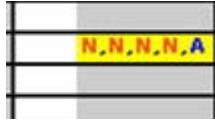
Figure 12. Multiple-Color Formatting Within One Cell

Note that some preset colors provided by name may not be the desired shade. The following website is helpful for choosing and refining color choices (http://www.colorschemer.com/online.html). In our code below, we use a combination of named colors and refined HEX values using this website.

Source Code to Set Colors

```
data new4;set new3;
length o1-&outmaxvar c1-&commamaxvar $35 callstring $500;
array codes(*) $ e1-&evtmaxvar.;
array outcomes(*) $ o1-&outmaxvar;
do i=1 to &evtmax;
select (codes(i));
  when (330)
    outcomes(i)=' ^S={foreground=blue} A';
  when (336)
    outcomes(i)=' ^S={foreground=#C6538C} B';
  when (491)
    outcomes(i)=' ^S={foreground=black} C';
…
  when (360,362)
    outcomes(i)=' ^S={foreground=#B8005C} R';
  otherwise ;
end;              /* end of select */
end;/* end of DO */
array outcomes2(*) $ o2-&outmaxvar;
array commas(*) $ c1-&commamaxvar;
do i=1 to &evtmax-1;
if outcomes2(i) ne '' then commas(i)=' ^S={foreground=black} ,';
end;
callstring=&callstr;
keep fidu callstring;
run;
```

When we rework the SAS Management Console report and focus on the formatting, we may come to the conclusion that less is more. Figure 13 contains the same paradata as the report in Figure 11 but with less-intensive formatting. The result is cleaner and allows the user to focus on the problem cells highlighted in yellow. Previously, the cyan blue columns for the weekend columns and the contrast between that and the lighter green columns for the weekend day competed visually with the highlighted yellow cells. For a report, the user needs to focus on the problems highlighted in yellow and then to notice the day of week. By simplifying—utilizing the principle of similarity to group like elements—we begin to arrive at a cleaner presentation that draws attention to those cells that require being noticed first.

Figure 13. SAS Management Console With Contrasting Colors for Weekend Times Eliminated

If we simplify even further and take away the colors for the text that describes the types of cases—non-contact, refusal, appointment, complete, final incomplete, other, and breakoff—the highlighted cells become even more visually dominant and important (Figure 14).

Figure 14. SAS Management Console With Simplified Header

9

Now, we can darken the weekend hours to make a visual group so that there is a division between daily and weekend activities. We notice the distinction, but it does not distract from the highlighted yellow cells (Figure 15).

Figure 15. SAS Management Console With Header Simplified and Weekend Columns Darkened

After each format iteration, it is important to take a step back and evaluate whether the change takes away from the visual effect or adds to it. In our example, we find that the color-coded result adds value to the visual effect and do not conflict with the message of the report. Once we reinstate the colored text, we find it easier to detect what is represented in each block of time and that it aids in drawing conclusions. Although we highlight situations we know are suspicious, we do not claim to identify all circumstances that need closer inspection. By seeing the different-colored events within shaded blocks of time, the analyst can easily understand how the FI is spending his or her time, and we have achieved our goal. We now bring back the colored letters associated with different events and assess the results. This version is clearer, but some letters are still getting lost, especially the green, and the white columns compete with the letters and highlighted cells for our attention (Figure 16).

Figure 16. SAS Management Console With Colored Events Reinstated

Next, we darken the cells behind the percentage columns N, R, and C to make a distinction between the percentage rate columns and the days of the week. We decrease the contrast between the "Day" and "Night" columns, thus focusing more attention on the colored letters and highlighted cells. We darken the cells just below "Day" so that the distinction between the "Night" and "Day" columns during the week does not get lost. Finally, we decide that the color associated with the letter C would work as a dark gray and that the red for N needs to be much darker in order to be truly visible (Figure 17).

CASEID	N	R	C	Mon Day	11-Aug Night	Tue Day	12-Aug Night	Wed Day	13-Aug Night	Thu Day	14-Aug Night	Fri Day	15-Aug Night	Sat Day	16-Aug Night	Sun Day	17-Aug Night
	79%	5%	64%														
	31	2	25														
1000001	3	0	0						N,N,N		N					N,N	
1000002	0	0	1						A								
1000003	4	0	1														C
1000004	2	0	1					C									
1000005	2	0	1						N,N,N,N,A					C			
1000006	5	0	0														
1000007	4	0	0														
1000008	0	0	1														
1000009	1	0	1						A,N							C	
1000010	2	0	1														
1000011	1	0	1						N,N,A					C			

Figure 17. SAS Management Console with reduced contrast

In the end, as much as we love to design sound and attractive reports, we must step back to reevaluate the applied formatting to ensure we have not overstepped the needs of the analyst.

We suggest the creation of a grayscale version of the report during the process of creating and finalizing. This process step offers a check on the data as well as a confirmation that formatted areas of color have not inadvertently created confusion rather than clarity (Figure 18).

CASEID	N	R	C	Mon Day	11-Aug Night	Tue Day	12-Aug Night	Wed Day	13-Aug Night	Thu Day	14-Aug Night	Fri Day	15-Aug Night	Sat Day	16-Aug Night	Sun Day	17-Aug Night
	79%	5%	64%														
	31	2	25														
1000001	3	0	0						N,N,N		N					N,N	
1000002	0	0	1						A								
1000003	4	0	1														C
1000004	2	0	1					C									
1000005	2	0	1						N,N,N,N,A					C			
1000006	5	0	0														
1000007	4	0	0														
1000008	0	0	1														
1000009	1	0	1						A,N							C	
1000010	2	0	1														
1000011	1	0	1						N,N,A					C			

Figure 18. SAS Management Console Grayscale Check

An example of a potential conflict introduced by color would be the choice of red to draw attention to a cell. This is not always a bad choice, but if introduced as a letter in the above report in one of the darker cells, that letter might disappear entirely in the grayscale version. Red is a much darker color than is often realized—not to mention any issues with color blindness, which is not covered in this paper.

Finally, it should be noted that even in our electronic-minded workplace, many people still rely on printouts and do not always have color printers available. Designing the report for an effective grayscale presentation will solve that issue.

CONCLUSION

While SAS provides the tools to customize extensively, it may be more effective to keep the formatting to a well-chosen minimum. It is important to design reports so that any visual additions add to the understanding of the data rather than the contrary. When too many layers of shading and colors are applied, we run the risk of overstimulating the eye and losing focus on key elements, information we want the user to see.

Our walking through the redesign of these two programs provides examples of how one can approach simplifying the output, thereby enhancing and supporting the intent of the report.

REFERENCES

- Swoope, Elizabeth A. "SAS ODS Style Examples." 2011. Available at http://stat.lsu.edu/SAS_ODS_styles/SAS_ODS_styles.htm.

- Tufte, Edward R. 1990. *Envisioning Information.* Cheshire, CT: Graphics Press.

RECOMMENDED READING

- "Principles of Grouping." *Wikipedia.* May 2012. Available at http://en.wikipedia.org/wiki/Principles_of_grouping.

- Spokane Falls Community College. "The Gestalt Principles." No date. Tutorial and Exercises. Available at http://graphicdesign.spokanefalls.edu/tutorials/process/gestaltprinciples/gestaltprinc.htm

- Todorovic, Dejan. "Gestalt Principles." *Scholarpedia.* 2008. Available at http://www.scholarpedia.org/article/Gestalt_principles

- Tufte, Edward R. 2006. *Beautiful Evidence.* Cheshire, CT: Graphics Press.

- Tufte, Edward R. 2001. *The Visual Display of Quantitative Information.* Cheshire, CT: Graphics Press.

- Tufte, Edward R. 1997. *Visual Explanations.* Cheshire, CT: Graphics Press, LLC.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Helen Smith
RTI International
3040 Cornwallis Road, P.O. Box 12194
Research Triangle Park, NC 27709-2194
Work Phone: 919-541-6912
Fax: 919-541-6178
E-mail: hsmith@rti.org

Susan Myers
RTI International
3040 Cornwallis Road, P.O. Box 12194
Research Triangle Park, NC 27709-2194
Work Phone: 919-541-7441
Fax: 919-990-8368
E-mail: smyers@rti.org

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.