A Unique Approach to Create Custom Reports by Leveraging the Strengths of SAS® and Excel®

Amy O. Wilkerson,* Barbara S. Bibb, Brett P. Anderson, Mai T. Nguyen • RTI International, Research Triangle Park, NC

More Information

***Presenting author:** Amy Overby Wilkerson
919.541.8021 | awilkerson@rti.org

RTI International
3040 East Cornwallis Road | Research Triangle Park, NC 27709

Presented at: SAS Global Forum 2013, city, ST, month 00–00, 2011

www.rti.org

Abstract

Survey projects often require custom reports to allow project staff to monitor production as well as produce various statistics from the collected data. At RTI, we've come up with a unique approach for creating custom reports for our projects by leveraging the strengths of SAS® and Excel®. In SAS® we use PROC SQL to select and when necessary aggregate data. After processing the data in SAS®, results are sent to Excel® for reporting and graphics.

1. Introduction

A client requested an Excel® report showing the most up-to-date information from an online survey that would be generated automatically overnight. The survey itself was programmed in Hatteras®, a proprietary software program developed by RTI International. The report used both data collected by the survey as well as data coded by the systems.

This process takes advantage of the features of Excel®, SQL Server®, and SAS®. A worksheet is built using the formulas required for the desired report; and then data is extracted using PROC SQL. This data is output to separate Excel® worksheets; the first worksheet, containing the formulas, picks up the data from the data storage worksheets for the final calculations and formatting.

An example of an Excel® report is shown in **Table 1** below. Two graph output examples are shown in **Figures 1 and 2**.

2. Final Excel Report

Table 1. Example Report in Excel

Overall	Count	% of Complete
Number Completed Patient Surveys	**201**	**100%**
Self interview (web)	76	62%
Interviewer (telephone)	125	38%
Completed by Gender	**Count**	**% of Complete**
Male	28	14%
Female	173	86%
Not answered	0	0%
Total	**201**	**100%**
Completed by Age	**Count**	**% of Complete**
40 years old or younger	5	2%
41 – 50 years old	12	6%
51 – 60 years old	60	30%
61 – 70 years old	67	33%
71 years old and older	57	28%
Not answered	0	0%
Total	**201**	**100%**
Completed by Time Taking Forteo	**Count**	**% of Complete**
Less than 2 months	25	12%
2 months to less than 4 months	53	26%
4 months to less than 6 months	40	20%
6 months to 1 year	43	21%
More than 1 year	40	20%
Not answered	0	0%
Total	**201**	**100%**

3. Steps to Create Excel Reports

Step 1. Set Up Lib Names and Blank Excel® Worksheets to Hold and Summarize Codes

```
libname in '\\...\Patient Dataset';
libname out '\\...\Reports';
libname REMS ODBC complete='DRIVER=SQL Server;
   SERVER=rtpwsql12;UID=xxx;PWD=xxx;DATABASE=xxx;';

libname xls excel
   '\\...\Patient_Survey_Response_Report.xlsx';
libname codes excel '\\...\patient_codes.xls';

%let sumstat_codes = 'sumstat_codes$'n;
%let gender_codes = 'gender_codes$'n;
%let age_codes = 'age_codes$'n;
%let time_codes = 'time_codes$'n;
%let mode_codes = 'mode_codes$'n;
```

Step 2. Add a Section to Re-Initialize the Data

```
* Remove existing data first;
proc datasets lib=xls nolist;
    delete sumstat_report
           gender_report
           age_report
           time_report
           mode_report;
run;
```

Step 3. Write PROC SQL Procedures to Categorize Each Data Point of Interest

```
* Categorize sumstat codes;
proc sql;
  create table xls.sumstat_report as
  select code, sum(sumstat_count)
  from temp
  group by code;
quit;
```

4. Conclusion

You can see the report tab displayed with formulas and the data storage tabs associated with the report in **Table 2** below.

Table 2. Formulas in Excel

	Overall	Count	% of Complete
7	**Number Completed Patient Surveys**	**=B8+B9**	**=C8+C9**
8	Self interview (web)	=mode_report!B3	=(mode_report!B2)/B7
9	Interviewer (telephone)	=mode_report!B2	=(mode_report!B3)/B7
10	**Completed by Gender**	**Count**	**% of Complete**
11	Male	=gender_report!B2	=(gender_report!B2)/B7
12	Female	=gender_report!B3	=(gender_report!B3)/B7
13	Not answered	=gender_report!B4	=(gender_report!B4)/B7
14	**Total**	**=B11+B12+B13**	**=C11+C12+C13**
15	**Completed by Age**	**Count**	**% of Complete**
16	40 years old or younger	=age_report!B2	=(age_report!B2)/B7
17	41 – 50 years old	=age_report!B3	=(age_report!B3)/B7
18	51 – 60 years old	=age_report!B4	=(age_report!B4)/B7
19	61 – 70 years old	=age_report!B5	=(age_report!B5)/B7
20	71 years old and older	=age_report!B6	=(age_report!B6)/B7
21	Not answered	=age_report!B7	=(age_report!B7)/B7
22	**Total**	**=B16+B17+B18+B19+B20+B21**	**=C16+C17+C18+C19+C20+C21**
23	**Completed by Time Taking Medication**	**Count**	**% of Complete**
24	Less than 2 months	=time_report!B2	=(time_report!B2)/B7
25	2 months to less than 4 months	=time_report!B3	=(time_report!B3)/B7
26	4 months to less than 6 months	=time_report!B4	=(time_report!B4)/B7
27	6 months to 1 year	=time_report!B5	=(time_report!B5)/B7
28	More than 1 year	=time_report!B6	=(time_report!B6)/B7
29	Not answered	=time_report!B7	=(time_report!B7/B7)
30	**Total**	**=B24+B25+B26+B27+B28+B29**	**=C24+C25+C26+C27+C28+C29**
31			

Report | age_report | time_report | sumstat_report | mode_report | gender_report

5. Discussion

The remainder of the SAS® program repeats the process just described for the other variables of interest, inserting and storing the data on the support worksheets. The final step is updating the formulas on the Patient Survey Response Report worksheet to refer to the counts of variables on the other worksheets.

6. Alternative Graphic Output

In addition to outputting data as an Excel spreadsheet, the output results can also be displayed as a graph:

Figure 1. Bar Graph in Excel

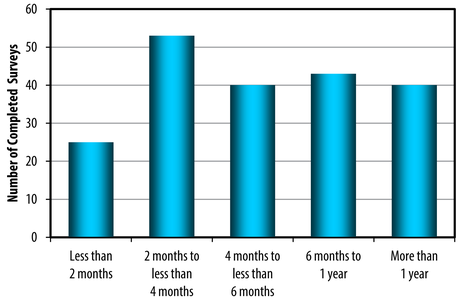

Figure 2. Pie Chart in Excel