ABSTRACT

Report Production System REPORT LINE is a program that makes it easy to produce custom reports from your SAS System data set, and it can be integrated to any your applications based on SAS Software.

REPORT LINE is a way for you to solve your application problems while maintaining a great degree of flexibility of data processing and data presentation.

REPORT LINE is based on relational data dictionary and preserves the data integrity by the powerful verification facility. While other reporting methods require a lot of programming, no programming is required with REPORT LINE.

This article is targeted at SAS System end-users and applications designers.

INTRODUCTION

What makes REPORT LINE unique?

REPORT LINE is not like other report writers. It's in a class by itself in these ways:

Code-free - To produce a new report without REPORT LINE a programmer would have to write a new program. With REPORT LINE, you no longer need to write programming instructions in order to produce a new report. REPORT LINE reports can be produced with only a few simple data definitions. These definitions not use any syntax, these simple data into the tables. Each table defines one or few aspects of the report you want (see Fig.1). For example, Reptitle table contains descriptions of the titles you want the report to have. The Repclass table contains definitions of the classification hierarchies. These tables form the data dictionary and you only tell REPORT LINE the result you want, not the details of how to produce it.

Quality of reports - the reports produced by REPORT LINE look every bit as professional as those produced by individual report programs. Column headings are neatly aligned over the data, titles are correctly justified, etc. This means you can use REPORT LINE to produce production reports. REPORT LINE supports "sheet breaks" (not just "page breaks") to facilitate distribution of report sections to different people.

Versatility - for those who want to control every last aspect of how the report looks, there are options available to do just that. You can control everything from where the date goes in a title, to the number of decimal places to show for a numeric field, to what the total line should say, etc. Powerful functions let you perform complex calculations.

In addition, REPORT LINE has the unique options that lets you hebrew/arabic custom reports that includes text right alignment and report right to left orientation.

REPORT LINE FEATURES

REPORT LINE provides powerful methods to create user-defined tabular reports, a great degree of flexibility in classification hierarchies, a variety of mechanisms for titling and formatting variables and statistics.

The following features make REPORT LINE powerful and flexible report creating tool:

- a great number of classifications hierarchies levels may be defined
- two types of classifications lets you split data to different tables, or groups of rows in the same table
- different statistics may be assigned to each analysis variable
- report may contain at the same time either raw data and calculated statistics
- full hebrew/arabic support includes sorting, report right to left orientation, text right alignment
- automatic dynamic page allocation
- data of report data dictionary are easy to read, easy learned by non-technical users
- validity checks data before processing it.

The statistics that REPORT LINE computes are many of the same statistics computed by descriptive procedures of SAS System.

WHO CAN USE REPORT LINE

REPORT LINE was designed with both applications designers and end-users in mind.

- Applications designers can use REPORT LINE for fast, easy-to modify production reports for any application testing or verifying new report's definitions integrated into any application based on SAS Software.
REPORT LINE lets end-users have the information they want without intervention from programmers.

HOW REPORT LINE WORKS

REPORT LINE runs under SAS System in the batch or on-line mode.

REPORT LINE input

There are three data sources for REPORT LINE input. First of them is the raw data set for processing. The second one are the tables of REPORT LINE data dictionary, which contains the report definition. The mentioned data sources are usual SAS System data sets. The third data source is the set of passing parameters to the REPORT LINE macro procedure (see below).

Return codes

Using three combined sources of the inputed information, the REPORT LINE executes, then issues one of the following return codes to describe the results of the run:

No errors were issued (return code is equal 0) or one or more severe errors were detected (return code is equal 1).

REPORT LINE output

The REPORT LINE produces two types of output:

• requested report (if return code equal 0)
• error messages report (if return code equals 1).

REPORT LINE produces an output sequential file in ASCII/EBCDIC format. The tables/subtables of the output report can contain more rows than can be printed on one physical page that defined by passing parameters. If so, REPORT LINE automatically divides the report's table/subtables into subsets and writes to output file. REPORT LINE always allocates each report's subtable on the separate physical page.

PRODUCING REPORTS

He, who at least once met with data analysis problem, knows, that different ways of sorting and classifications discover new data features. That’s why definition of different reports on the same data - is the powerful method to receive many-sided information. Easiness and simplicity of report definition lets you define different points of view on data.

REPORT LINE has a very short learning curve, because the only thing you need to know is the order of the REPORT LINE data dictionary tables filling.

REPORT LINE data dictionary

To define a report means to fill data in the tables of REPORT LINE data dictionary. Each of the tables is responsible for the certain features of the report (Fig.1). Once the report was defined, the definition will be stored as a data in the SAS data sets, and the report may be produced each time you need, by referencing to the stored data. The tables of REPORT LINE data dictionary are as follows:

1. The report declaration table Reptitle contains general report title, name of raw data set, general report attributes. The report identical number for each report is defined in this table the very first time, and that’s why this table is required to be filled in for every report definition.

2. The report subtitles declaration table Reptable contains a list of report subtitles and references to the common use parameters (table Repparm), which will be described further.

3. The report variables declaration table Rephead contains variables names from the raw data set and definitions of their attributes and formats.

4. The report classification table Repclass contains the types of classifications and list of classifications variables.

5. The report statistics definition table Repsum contains the keywords of requested statistics associated with analysis variables and classification, if any. It means that statistics, which are defined by keywords are calculated for the certain analysis variable within the group received by classification, or over all observations if there is no classification defined. The titles for statistics are also defined here.

6. The common use parameters table Repparm contains the names and the values of user-defined parameters, which are available for all reports defined in the data dictionary.

7. The variables legends table Replegen contains the codes (short values) and corresponding meanings (extended values) of variables values from raw data set.

REPORT LINE options

A great degree of REPORT LINE flexibility is provided by set of the following options:
• **rattr** field of Reptitle table, that defines the form of the report and may be one of the followings:
  - **S** - such value prescribes to print the report in a short form, without raw data; such form is very effective when you have created the report on large amount of data
  - **C** - prescribes the condensed form of the report, when it contains raw data and computed statistics, but there is no lines that divide data to groups
  - **missing value** - presents either raw data and computed statistics including lines that divide data to groups.
• **rattrib** field of Rephead table, that defines actions on the variables and may be combination of the followings:
  - **S** - prescribes to sort raw data by the variable in ascending order
  - **L** - this value means, that the values of the variables should be changed to their legend, defined in Replegen table.
  Note: each variable with "L" value of rattrib receives character type, so you have to care about format correspondence, and you can associate only those statistics available for the analysis variable measured on a nominal scale
  - **T** - prescribes to treat the variable values as a text, so this variable will receive a special processing according to language
  - **missing value** - means formatted printing of the variable values by the default.
• **rtype** - the field of Repclass table, that defines type of classification and may be one of the followings:
  - **T** - means that the variable will be used as a table classifier, a change of the variable's value leads to creating of the next subtable of the output report
  - **R** - defines the variable as row classifier, a change of the variable's value leads to creating of the next group in the same table/subtable
  Note: each variable, which was defined as a table or row classifier must receive "S" value of rattrib option (see above).

**REPORT LINE data dictionary verification**

REPORT LINE provides facility for the verification of a report definition. This facility checks data dictionary according to the report definition rules. In the case of non correct report definition REPORT LINE creates the special report with the error messages and possible sources of the error. The verification also checks that the raw data set structure corresponds to the report definition. The verification facility is optional. It's recommended to use it for the first time producing of each report (especially for the beginners).

**How to run REPORT LINE Macro**

In the SAS System terms REPORT LINE is a macro procedure. It can be used either as the part of different SAS System applications or in stand alone mode. To call REPORT LINE in stand alone mode the next steps must be executed:
1. definition of the necessary library references
2. including of REPORT LINE macro procedure into SAS
3. calling the procedure.

Usage of the macro procedure REPORT LINE is the following:

```
%repline(repid, librep, libdat, fname, lang, psize, isize, misval, fiver, libver) ;
```

where these parameters as:

- **repid** report identification number
- **librep** library reference to the REPORT LINE data dictionary
- **libdat** library reference to the input raw data set
- **fname** full filename for the report output
- **lang** language declaration, it must be L (left to right report orientation) or R (right to left report orientation)
- **psize** specifies the number of lines that can be printed per page, the value can be from 20 to 500
- **isize** specifies the printer line width for the print file, it's values can range from 64 to 256
- **misval** specifies the character that defined as missing value for all variables
- **fiver** verification flag, 1 prescribes to make and 0 to pass verification process
- **libver** library reference to the REPORT LINE verifications temporary data sets, it must be already defined if parameter fiver equal 1. libver it's different from librep and libdat libraries references

When REPORT LINE macro procedure has finished it's run you can use the %REPRC macro function to determine it's return code. The syntax is:

```
rc = %reprc ;
```
Keywords and formulas

Below is a list of the keywords and corresponding statistics that can be specified in the table Repsum:

- **N** - Number of observation in the group with non missing values
- **A** - Average
- **D** - standard Deviation
- **X** - Maximum value
- **S** - Sum
- **V** - Variance
- **R** - Range
- **P** - Percent of occurrence
- **Q** - percent of Quantity.

All these statistics are available for the analysis variable measured on a continuous scale. Only N and P statistics are available for the analysis variable measured on a nominal scale.

Missing values

Missing values may occur either in class and analysis variables. While in the class variable, missing values are considered as valid levels. Each statistic associated with an analysis variable is calculated over nonmissing values. The misval parameter of REPLINE macro, helps to integrate the REPORT LINE into another SAS System based application with it's missing code.

CONCLUSION

REPORT LINE implements the table-driven technology and has a completely open architecture based on industry standards, which allows the product to be completely portable. REPORT LINE processes huge data tables and can be integrated to any SAS System applications. REPORT LINE was integrated in the code-free, SAS Software based, application production system TARGET, developed by the authors [1].

The REPORT LINE software is available from the authors, which may be reached at the following address:

Samuel Berestizhevsky and Tanya Kolosova
P.O. Box 3294, Nazareth-Elit 17000, Israel
Fax/phone (972) 6-554254

REFERENCES


SAS is registered trademark of SAS Institute Inc. in the USA and other countries.