

SAS Dictionary Tables zur Nutzung von Selektionen am Beispiel des AXA-CRM-Explorers

Thomas Rüdiger
AXA Service AG
Marketing/Direktvertrieb
50828 Köln (INK-4.49)
thomas.ruediger@gmx.at

Zusammenfassung

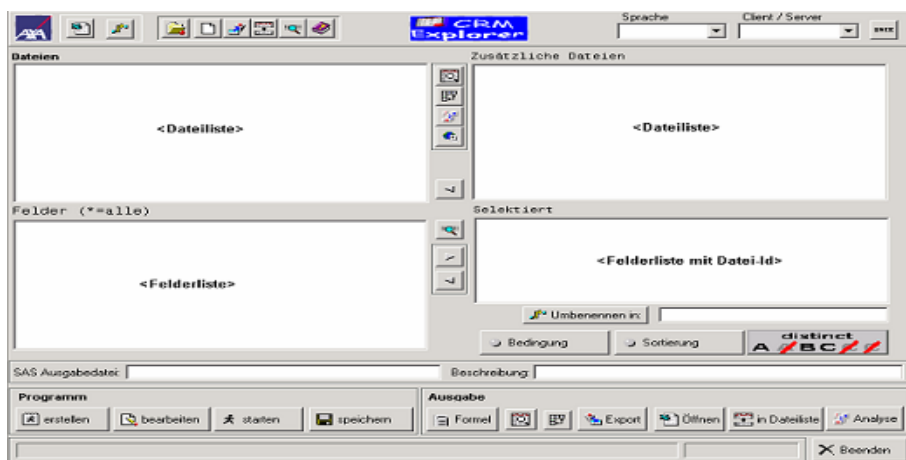
Neben SAS-Spezialisten gibt es in Unternehmen Anwender mit hohem Selektions- und Analyse-Bedarf. Statt zeitintensiver Anfragen an Analytiker besteht ein allgemeiner Wunsch an "Quick and Clean"-Lösungen. Das im AXA-Kampagnenmanager eingesetzte Business Objects erwies sich als inflexibel hinsichtlich neuer Datenstrukturen bzw. neuer Kampagnendaten. Neben hohem administrativen Aufwand fielen zusätzliche externe Kosten an (BO-Designer). Die AXA-Lösung beruht auf einer SAS/AF-Oberfläche, die die SAS-Dictionary-Tables nutzt ('CRM-Explorer').

Keywords: CRM Explorer, SAS/AF, SAS Dictionary Tables, Referential Constraints, PROC CONTENTS, PROC MEANS, PROC FORMAT, PLZ-Graphik.

1 Die Oberfläche des Data Managers innerhalb des CRM Explorers

Der CRM-Explorer ist eine komplette SAS/V8-Lösung und startet mit einer Auswahlliste von CRM-Highlights.

Der 'Data Manager' im CRM Explorer erstellt und startet das SAS-Selektions-Programm. Bedingungen und Sortierung sind über Untermenüs regulierbar.



Der CRM-Explorer (derzeit in SAS/V8) benötigt nur wenige SAS-Module. Auf der SAS-Lizenzseite entstehen keine zusätzlichen Kosten außer den ‘Standard-Modulen‘:

- SAS/BASE und SAS/STAT
- SAS/ACC-PC File Formats und/oder SAS/ACC-ODBC
- Datenbank-Schnittstellen nach Bedarf (DB2, Oracle, ..)
- SAS/AF (nur für Weiterentwicklung)

2 SAS Dictionary Tables

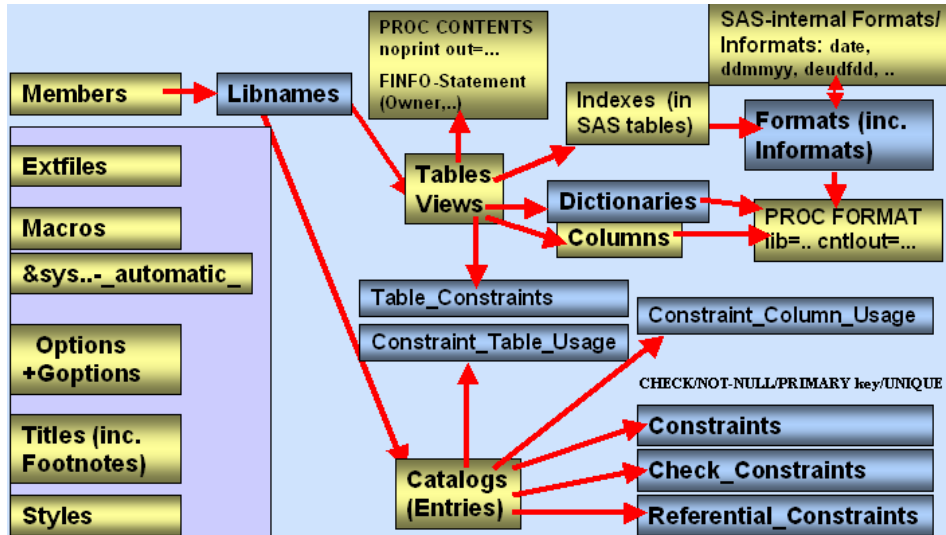
SAS-Dictionary Tables sind spezielle Read-Only-Data Views, die Informationen über den aktuellen Stand einer SAS-Session beinhalten und nur über SQL-Syntax oder über vorbereitete Views in SASHELP erreichbar sind: z.B. ist SASHELP.VCOLUMN ein SQL-View auf DICTIONARY.COLUMNS.

Einige Prozeduren wie PROC DATASETS, PROC CONTENTS, PROC FORMAT und einige Elemente des Data Steps {attrn(DSID, ‘*’) mit *=nobs, nvar, ..} bedienen sich direkt der Dictionary Tables.

Permanente SAS-Dateien ohne Libname-Bezug in der neuen Quote-Notation sind über die Dictionary Tables nicht erreichbar:

```
data `c:\SAS_Datei`;Test=1;run;
```

Nähere Informationen zu Dictionary Tables sind im Internet zu finden unter:
<http://v9doc.sas.com/cgi-bin/sasdoc/cgigdoc?file=../proc.hlp/a000146905.htm>



3 User-Schichten mit &sysuserid.

Der CRM-Explorer kennt derzeit über automatische Client-SAS-Makrovariable &sysuserid. zwei User-Schichten

- a) SAS-Administratoren für SAS/AF-Design und Tabellenbereitstellung und
- b) Anwender mit den Schwerpunkten: Auswahl von Feldern, Bedingungen, Sortierungen, Mengenlehre (Join Types), distinct/non-distinct-Entscheidung

4 Dateiauswahl

Die Input-Dateien werden vom Administrator des CRM-Explorers über eine Liste mit verfügbaren Dateien in den CRM-Explorer integriert - Vorteil: geringer Administrationsaufwand! (Neue Datei=eine Zeile in Liste)

Zusätzliche externe Daten (xls, mdb, dbf, sas7bdat) werden mit zwischengeschalteter Qualitätsprüfung eingelesen. Der Datenimport beinhaltet folgende Qualitätsprüfungen:

T. Rüdiger

Identifikation xls-Arbeitsmappen, mdb-Tabellen über ODBC und über die _library-Methode in SAS/AF.

```
%let DATA=%str(c:\daten.xls);
%let Path=%str(DSN=Excel-Dateien;DBQ=&DATA.);
libname MSEXCEL odbc required="&Path.";
data MSEXCEL.MSEXCEL2;
set MSEXCEL.'Datenreihe_Basis$'n;
run;
```

Änderung von Variablenamen in V8-Notation

```
options validvarname=any; ...
Rename '1. Übung'n =UEBUNG
```

Typ-Transformation, z.B. put(PLZ,z5.) als PLZ

Existenzprüfung Key-Variablen (PNR, ...)

Prüfung des Primary Keys auf Eindeutigkeit

Import ausschließlich in benutzereigene Library analog SASUSER-Library zur Vermeidung von ungewollten Updates

Datei wird an Datei-Auswahlliste des CRM-Explorers angehängt

5 Verknüpfungsregeln mit PROC CONTENTS

Der CRM-Explorer präferiert die PROC CONTENTS zum Auslesen der Variablenattribute (Table-Header)

```
* CONTENTS BY PROC CONTENTS;
proc contents noprint data=PARENT out=mytable;
run;
* CONTENTS BY SQL;
proc sql; create table mytable as
  select * from dictionary.columns
  where libname='WORK' and memname='PARENT';
quit;
* CONTENTS BY SASHELP;
data mytable;
  set sashelp.vcolumn(
    where=(libname='WORK' and memname='PARENT'));
run;
* CONTENTS BY OPEN/CLOSE-Statement;
data mydata;
  RC=open('PARENT');
```

```
do i=1 to attrn(RC, 'varnum');
  ...;
end;
RC=close(RC);
run;
* CONTENTS BY SCL-List; *siehe: _dictionary-Methode;
```

Vorteile von PROC CONTENTS: leichteste Notation+ i.d.R. schnellster Zugriff (abhängig von Anzahl verfügbarer Tabellen/Views)

Join-Variablen (Primary/Foreign Keys) zwischen je zwei Tabellen werden aus dem PROC CONTENTS über Feldnamen und Hierarchiestufen (Granularität) definiert

Im Unterschied zu SAS/V9, bei der generelle Key-Beziehungen zwischen Parent und Child Table über Referential Keys definierbar sind, werden im CRM-Explorer Key-Beziehungen dynamisch mit mehrfachen PROC CONTENTS definiert!

Frei definierbare Join-Beziehungen

(z.B. 'on substr(a.Organummer,1,4)=b.RBLNR')

sowie andere Verknüpfungen als zwischen Parent und Child Table oder 'Brother-Sister-Tables' werden demnächst genauso realisiert wie die Einbindung von Interessentendaten über Deduplikation.

Beispiel-Dateien: Der CRM Explorer erkennt über das PROC CONTENTS die gemeinsame Variable Sales_Region der Tabellen PARENT und CHILD als Verknüpfungsvariable

```
* Beispieldatei Parent-Table
  PARENT (Tagesumsatz pro Region);
data PARENT (keep=Date Sales_Region Sales);
Date=date(); format Date deudfdd10.;
do Sales_Region =1 to 10;
Sales=round(100000*ranuni(date()),.2);output;end;
label Sales_Region='Region' Sales='Umsatz'; format Sales
commax9.2;run;
```

```
* Beispielformat 'Region';
proc format lib=WORK.FORMATS;value Region 1='RBL
Sandler' ...;run;
proc catalog c=WORK.FORMATS;modify Region.format
(description='RBL');quit;
```

T. Rüdiger

```
proc datasets lib=WORK;modify PARENT;format Sales_Region
region.;quit;

* Beispiel Child-Table CHILD (Letzter Umsatz pro
  Region und Kunde);
data Child(keep=Sales_Region PNR Sales_PNR);set
SAMPLE;Weight1=0;
do PNR =10*Sales_Region to 10*Sales_Region+3;Weight1+1;
if Weight1<=2 then Weight=2**Weight1;else Weight=8;
Sales_PNR=Sales/Weight;output;end;
label PNR='Kunde' Sales_PNR='Umsatz pro Kunde';format
Sales_PNR commax9.2;
run;
```

6 Datensicht mit PLZ-Graphik

Für Input- und Output-Tabellen bietet der CRM-Explorer statt dem SAS/BASE-viewtable und dem antiken SAS/FSP-fsview einen AF-interner TableView mit Zeilen-Drill-Down-Funktionalität und PLZ-Graphik-Feature

```
data _null_;
file_in=" http://routenplaner24.telemap.de
/<image>.jpg";file_out="c:\<image>.jpg";
infile in URL filevar=file_in lrecl=1 recfm=S nbyte=read
end=EOF_URL
debug proxy="<your-proxy-server" user="<your-id>"
pass="<your-password>";
file out filevar=file_out lrecl=1 recfm=f;
do while (not EOF_URL); input buffer $CHAR1. @@; if
_error_ then leave;
put buffer $CHAR1. @@; end;stop;run;
```

Datums-/Zeit- und benutzerdefinierte Formate werden im CRM-Explorer mit PROC FORMAT für die Weiterverarbeitung (z.B. Anzeige Wertelisten) in eine Ausgabedatei geschrieben.

Datums-/Zeitformate:

```
proc contents noprint data=PARENT out=mytable(
  keep=Name Format);
run;
data Date; set mytable(where=
  (Format in ('DDMMYY' 'MMDDYY' 'DATE' 'Yymm') or
  substr(Format,4,4)='DFDD'));
```

```
run;
* analog Time ...;
```

Benutzerdefinierte Formate:

```
* VALUE DESCRIPTIONS BY PROC FORMAT;

proc format lib=WORK.FORMATS cntlout=mytable;
  select REGION;
run;
```

Tabelle 1: Interaktive Metadaten des CRM Explorers

Metadaten	=>SAS-Code	=>Projektanlage
Tabellen/Felder Bedingungen Sortierungen Mengenlehre (Join Types) distinct/non-distinct- Entscheidung	Parallel zur Code- Erstellung erfolgt ein Syntax-Check (Anzahl Klammern und Anführungs- zeichen) und ein Verknüpfbarkeits- Test. Bei Auslassen von Keyvariablen aus zwischen- liegenden Hierar- chiestufen erfolgt eine Fehler- meldung.	Die zur SAS- Code-Erstellung dienenden Meta- daten lassen sich als 'Projektdatei' ablegen und zur Befüllung von SAS/AF-Objekten wiederverwenden.
Interaktive Messung Programmdurchlaufdauer Remote-Work-Library- Check: %sysfunc(libref(RWORK))=0		

7 Interaktive Metadaten

Aus den Anwenderanforderungen wird ein SAS-SQL-Code in einem Anwender-eigenen SAS-Catalog abgelegt. Als SAS/AF-Anwendung mit integrierter Batchverarbeitung sind im CRM-Explorer zwei Prozesse zu kontrollieren:

1. SAS/AF-Prozesse: Für Updates von AF-Frames aus der AF-Entwicklungs- in die AF-Produktiv-Umgebung werden Anwender-Logins/Logouts in einer Registrierungs-Tabelle mitprotokolliert
2. Batch-Prozesse: Sollte der Batchlauf ein Zeitlimit (z.B. 2 Min.) überschreiten, erscheint ein akustisches Signal bzw. eine

T. Rüdiger

```
zusammengesetzte Melodie
data _null_;
call sound(<Tonhöhe>, (1/<Dauer>)*300);
run;
```

8 Mehrdimensionale deskriptive Analysen

Mehrdimensionale deskriptive Analysen nutzen die CLASS-Komponente der SAS-Prozedur PROC MEANS

```
proc means noprint | proc summary data=<Datei>;
class <Variablen>;
var <Variablen>;
output out=<Datei>
      n= mean= sum= std= min= median= max= nmiss=;
run;
```

Der Analyse-Frame ist die vorbereitende Schaltzentrale für Excel-/ RTF-/HTML-Reports, graphischen und geographischen Output oder dient zur Weiterverarbeitung von Zwischenergebnissen bei vorhandener Key-Variable.

9 Graphische Auswertungen

Der CRM-Explorer nutzt die graphischen Möglichkeiten aus SAS/AF (Bar-, Box-, Line-, Area-, Pie-Charts, Join- Scatter- Surface-Plots)
Die Output-Graphiken lassen sich als gif-Datei in Reports einbetten.
Geoanalysen (GIS) erfolgen mithilfe PLZ-Kreisgemeinde-SchlüsselungAF-Maps lassen sich ebenfalls über die Snapshot-Methode als gif-Datei exportieren.

10 Zusammenfassung

Der CRM-Explorer verknüpft die Stärken des SAS-Systems mit den SAS-Dictionary Tables und benötigt nur wenige SAS-Module. Als SAS/AF-Applikation orientiert sich der CRM-Explorer flexibel an individuellen Anwender-Bedürfnissen. Anwender finden sich auch ohne Training schnell zurecht, positiver Nebeneffekt: SAS-Spezialisten haben mehr 'Mining-Time'.