Data2Office: Eine neue Methode der SAS-Datenpräsentation in MS-Excel

Ralf Neumüller, beyondbytes, 64372 Ober-Ramstadt
Stephan Frenzel, Kybeidos, 69120 Heidelberg

Trotz all der hervorragenden Möglichkeiten zur Datenpräsentation mit SAS (PROC REPORT, etc.) werden Berichte (Reports) von Fachabteilungen zumeist ein Form eines Excel-Sheets gewünscht. Zur Übertragung der Daten aus einem SAS Data set nach MS-Excel stehen mehrere Methoden bereit:

SAS/Base-Methoden zum unformatierten Export nach MS-Excel

- Excel Libname Engine
- PROC EXPORT
- PROC PRINTTO
- Data Step mit FILE und PUT
- ODS CSV

SAS/Base-Methoden zum formatierten Export nach MS-Excel

- ODS HTML,
- ODS Tagsets.Msoffice2k
- ODS Tagsets.ExcelXp (über XML)

Methoden die Microsoft-Technologien verwenden

- DDE
- ODBC

SAS-BI-Technologien

- SAS Add in for Excel

SAS/AF

Java Technologien über JDBC

- Actuate eSpreadsheet (ehem. Formular One)
- Apache Jakarta POI Projekt

Literatur zu diesen Technologien liegt mannigfaltig im Netz vor. Einfaches Googlen (export sas data to excel), ein Blick in das SAS-Wiki oder das SAS-Anwenderhandbuch der Uni Heidelberg helfen hier weiter. Eine kleine Literatursammlung befindet sich am Ende dieses Artikels.

Data2Office

All diese Methoden bieten jedoch nicht - oder nur sehr beschränkt - die Möglichkeit mit einfachen SAS/Base Programmen vorformatierte Excel-Templates zu befüllen. Diese Lücke möchte die neue native SAS Prozedur Data2Office (PROC D2O) schließen. Sie bietet eine neue Technik, um aus SAS heraus Excel-Arbeitsmappen zu generieren.
Das Konzept von Data2Office für SAS besteht darin, aus Excel-Vorlagen per SQL auf SAS-Dateien zuzugreifen - und das ohne den Umweg über SAS-Server (z.B. Share-Server). Der Zugriff erfolgt daher ausschließlich auf der Basis von Desktop-/Client-Funktionen.
Data2Office läuft unter allen Windows NT basierten Betriebssystemen (Windows NT, 2000, XP, Vista, 7) und ist kompatibel mit SAS 8.2 und höher sowie MS-Excel 97 und höher.

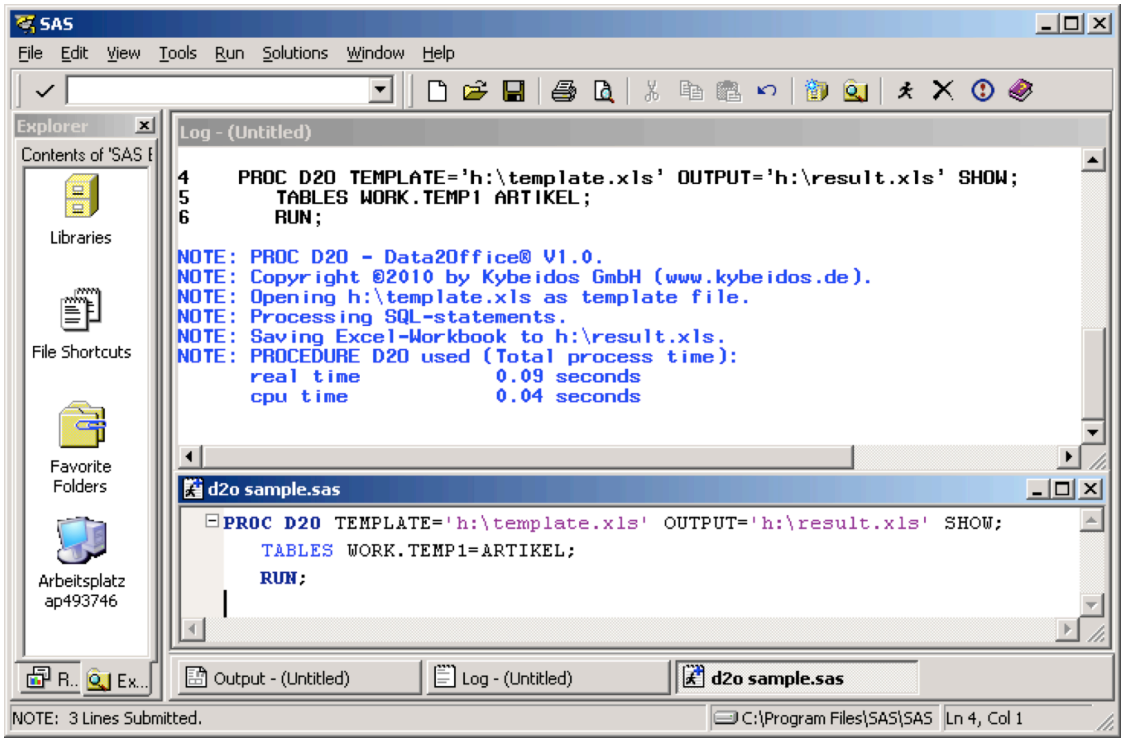
Der Data2Office-Kern ist schon seit Jahren in einer Reihe von Projekten erprobt. Dessen Implementierung in eine SAS Prozedur übertrifft bestehende Lösungen durch folgende Vorzüge:

- Hohe Performance
  (Massenbericht-Erzeugung durch Batch-Verarbeitung)
- Weitreichende Funktionalität
  (SAS-Datenimport in Grafiken, Formeln und VBA-Makros enthaltende formatierte Excel-Arbeitsmappen)
- Einfache Bedienung (native SAS-Prozedur)
- Schnelle Implementierung (SAS-Base und SQL Kenntnisse ausreichend)

Funktionsweise
Der Aufruf der SAS-Prozedur D2O ...

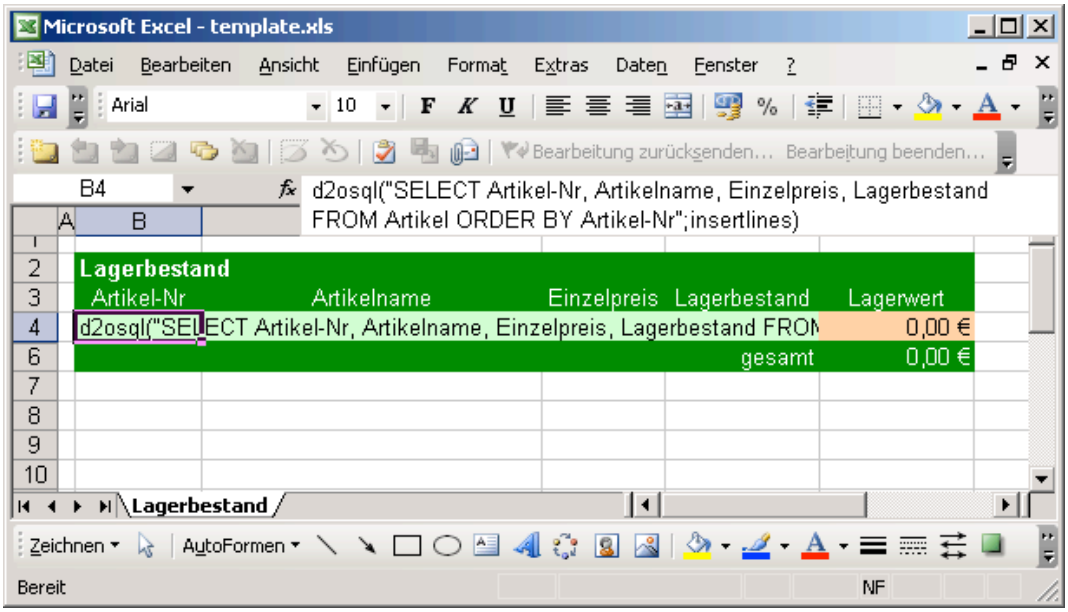
```
PROC D2O TEMPLATE='h:\Template.xls'
         OUTPUT='h:\results.xls'
         SHOW;
         TABLES WORK.TEMP1 = ARTIKEL;
RUN;
```

... stellt die SAS-Datei WORK.TEMP1 unter dem Namen ARTIKEL für den Zugriff aus Excel bereit und ruft die D2O-Engine zur Verarbeitung der Excel-Vorlage Template.xls auf.

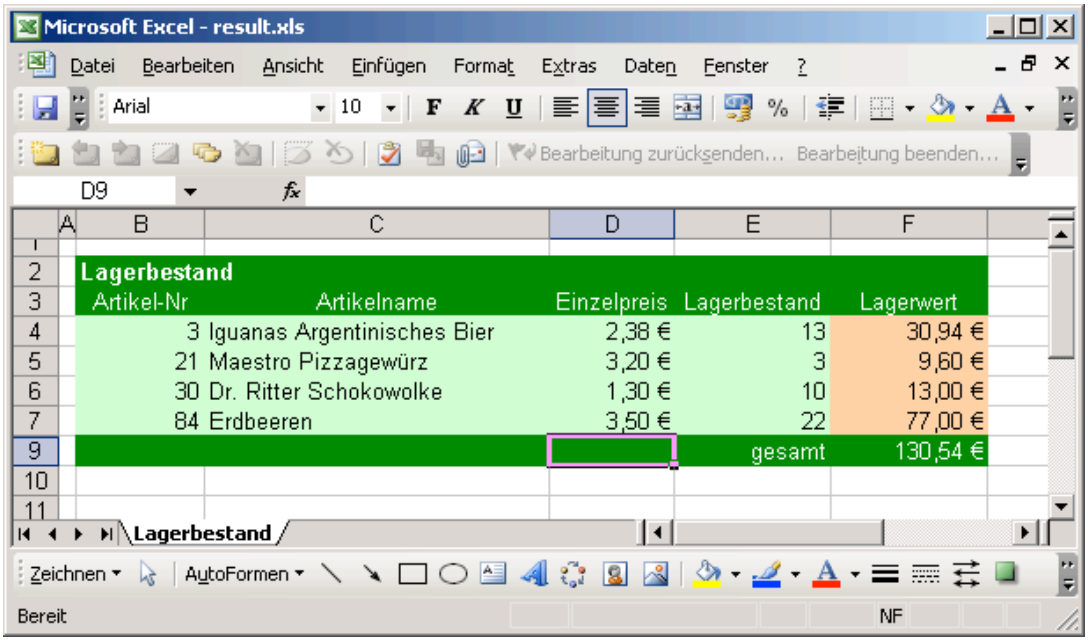


Mit der Funktion D2OSQL in der Excel-Vorlage werden die SAS-Daten dann selektiv nach Excel übernommen:

D2OSQL("SELECT Artikel-Nr, Artikelname, Einzelpreis, Lagerbestand FROM ARTIKEL ORDERBY Artikel-Nr.";insertlines)



Als Ergebnis wird das Excel-Sheet results.xls erzeugt, das die in D2OSQL angegebenen Daten des SAS-Data Sets WORK.TEMP1 enthält.



Die Funktion D2OSQL kann an mehreren Stellen in einem Tabellenblatt und auf verschiedenen Tabellenblättern in einem Excel-Dokument eingefügt werden. Die SQL-Befehle können sich dabei auf unterschiedliche SAS-Dateien oder beliebige Verknüpfungen (Joins) von SAS-Dateien beziehen. So können mit **einem** Aufruf von PROC D2O **mehrere** SAS-Dateien in Excel übernommen werden.

Bei der Datenübernahme bleiben in Excel berechnete Spalten innerhalb des Bereichs, in den Daten übernommen werden sollen, erhalten; sie werden beim Einfügen der Spalten übersprungen. Dadurch ist es möglich, in eingefügten Tabellen auch in MS-Excel berechnete Spalten zu verwenden.

Zur Steuerung des Einfügens von Zeilen und Spalten stehen verschieden Optionen der Excel-Funktion D2OSQL zur Verfügung:

D2OSQL (<SQL-Befehl>, insertlines)
→ komplette horizontale Zeilen werden eingefügt (empfohlen).

D2OSQL (<SQL-Befehl>, insertcells):
→ Es wird nur der notwendige Bereich eingefügt (falls rechts neben der Tabelle noch eine Tabelle ist).

D2OSQL (<SQL-Befehl>, noinsert):
→ Das Ergebnis überschreibt die vorhandenen Zellen. (Wenn man keine Summen-Zeilen hat, oder einen festgelegten Bereich z.B. für ein Diagram befüllt).

In den Excel-Vorlagen sind SQL-Funktionalitäten im vollen Umfang der Microsoft Jet SQL Engine nutzbar.

Nach dem Datenimport kann für die Nachbearbeitung der Excel-Arbeitsmappe ein VBA-Makro mit dem vordefinierten Namen D2OOPEN hinzugefügt werden. Dieses Makro wird - wenn vorhanden - vor dem Speichern der fertigen Excel-Arbeitsmappe ausgeführt. Damit hat man letztendlich die Möglichkeit den wirklich vollen Umfang der Excel-Funktionalitäten nach dem Befüllen mit Daten zu nutzen.

Perspektiven
Data2Office wird fortlaufend weiterentwickelt. Auf der Roadmap für zukünftige Versionen von PROC D2O stehen bereits eine Reihe von Funktionen, darunter die Generierung von PDF-Dokumenten und HTML-Seiten sowie die Unterstützung von Unix-Servern.

Lizenzmodelle
Folgende Lizenzstaffelung wird zur Verfügung gestellt:

- Test-Lizenz
  (90 Tage begrenzt, voller Funktionsumfang) kostenlos
- Edu-Lizenz
  (unbegrenzt, voller Funktionsumfang) 149,- €
- Workstation-Lizenz
  (unbegrenzt) 790,- €
- Server-Lizenz
  (unbegrenzt) 2490,- €

Download der Test-Lizenz
Eine Test-Lizenz von Data2Office steht auf der Website der Kybeidos GmbH bereit. Diese enthält folgende Dateien:

Setup_ProcD2O.exe – Installer-Datei mit D2O-Prozedur und –Engine
d2o-Beispiel.sas – Beispielprogramm mit D2O-Aufruf
grp_template.xls – Beispiel-Excel-Template zum Befüllen mit SAS-Daten

Ein ausführliches Benutzerhandbuch mit detaillierten Beschreibungen der Syntax und Beispielen ist in Bearbeitung.

Hersteller
Die "Data2Office"-Prozedur "D2O" ist ein Gemeinschaftsprojekt von Kybeidos (Konzeption und SAS-Prozedur, http://www.kybeidos.de) und Ralph Neumüller (Data2Office-Kern, http://www.beyondbytes.de).

Literatur
http://www.lexjansen.com/pharmasug/2006/posters/po13.pdf
http://rfd.uoregon.edu/files/rfd/StatisticalResources/data_trnsf2.txt
http://www.nesug.org/proceedings/nesug04/io/io09.pdf

http://www.urz.uni-heidelberg.de/imperia/md/content/urz/programme/statistik/sas-treff/2006-07-07.pdf
http://www.sas.com/offices/europe/austria/html/events/sas_club/downloads/sas_club19_bernadette.pdf
http://www.sas-consultant.com/professional/SEUGI18-Using-DDE-to-Pour-S.pdf
http://www.urz.uni-heidelberg.de/statistik/sas-ah/01.01.03/odbc/odbc.html
http://rfd.uoregon.edu/files/rfd/StatisticalResources/data_trnsf2.txt
http://www2.sas.com/proceedings/forum2007/168-2007.pdf
http://support.sas.com/rnd/papers/sugi29/ExcelXML.pdf