

Datenaustausch zwischen SAS 9.1.3 SP1 und MS Excel / Access 2003 mittels ODBC

Paul Eberhard Rudolph
Forschungsinstitut für die Biologie
landwirtschaftlicher Nutztiere (FBN)
Wilhelm-Stahl-Allee 2
18196 Dummerstorf
pe.rudolph@fbn-dummerstorf.de

Armin Tuchscherer
Forschungsinstitut für die Biologie
landwirtschaftlicher Nutztiere (FBN)
Wilhelm-Stahl-Allee 2
18196 Dummerstorf
atuchsch@fbn-dummerstorf.de

Carina Ortseifen
Universität Heidelberg, Universitätsrechenzentrum
Im Neuenheimer Feld 293
69120 Heidelberg
Carina.Ortseifen@urz.uni-heidelberg.de

Zusammenfassung

Die ODBC(Open DataBase Connectivity)-Technologie wurde von Microsoft entwickelt und stellt eine gemeinsame Schnittstelle für den Zugriff auf heterogene Daten zur Verfügung. Damit ist ein Austausch von Daten zwischen verschiedenen Anwendungen möglich. Die Grundlagen von ODBC und der Datenaustausch zwischen SAS (Version 6 bzw. 8) und MS Office-Komponenten werden außer in den entsprechenden SAS-Dokumentationen z. B. auch in Rudolph u. a. ([2], [3]) beschrieben.

In diesem Beitrag wird der Datenaustausch zwischen den aktuellen Versionen SAS 9.1.3 (SP 1) und MS Access und Excel 2003 beschrieben.

Grundlage der entsprechenden Darstellungen ist das Betriebssystem Windows XP.

Schlüsselwörter: ODBC, Datenaustausch, SAS 9.1.3 (SP 1), MS Access 2003, MS Excel 2003

1 Einleitung

Obwohl es inzwischen für fast alle gängigen Datenbanksysteme eigene SAS-Module zum Datenaustausch gibt, hat der Datenaustausch zwischen SAS und anderen Windows-Applikationen mittels ODBC (Open Database Connectivity) weiterhin seine Existenzberechtigung. Eine ausführliche Einführung in diese Technik ist außer in den entsprechenden aktuellen SAS-Dokumentationen [4], [5] auch in [1], [2] und [3] zu finden.

Um über ODBC Daten auszutauschen, werden vier Komponenten benötigt:

1. ODBC-fähige Anwendungen
2. ODBC-Datenquellen-Administrator
3. ODBC-Treiber
4. ODBC-Datenquellen (Data Sources).

In SAS wird der Import von ODBC-Datenquellen anderer Applikationen nach der Installation von SAS/Access to ODBC möglich. SAS-ODBC-Datenquellen sind jedoch erst nach der Installation des entsprechenden SAS-ODBC-Treibers definierbar, d. h., erst dann können SAS-Daten in andere Applikationen importiert werden.

In Kapitel 2 werden die Installation des aktuellen SAS-ODBC-Treibers und die Definition von SAS 9.1.3-ODBC-Datenquellen erläutert. In den Kapiteln 3 und 4 wird dann an Beispielen der Datenaustausch zwischen SAS (Version 9.1.3, SP 1) und MS Excel / Access 2003 mittels ODBC beschrieben.

Die beschriebenen Datentransfers wurden mit einer englischsprachig installierten SAS-Version 9.1.3 (SP 1) (Mindestumfang: SAS/CORE, SAS/BASE, SAS/Access to ODBC) unter Microsoft Windows XP Professional (SP 2, deutsch) realisiert. Unter Windows 98 oder Windows 2000 können die dargestellten Bildschirm- bzw. Fensterinhalte von den hier wiedergegebenen abweichen.

In den beschriebenen Vorgehensweisen sind Fenstertitel durch Unterstreichen/fett und die Auswahl aus in diesen Fenstern vorhandenen Schaltflächen oder Menüs durch Kursivschrift/fett gekennzeichnet.

2 SAS-ODBC-Datenquellen

Die Definition einer SAS(9.1.3)-ODBC-Datenquelle setzt die Installation eines SAS-ODBC-Treibers voraus. Ein derartiger Treiber ist auf der CD SAS Client-Side Components Volume 1 des SAS Installation Kit zu finden und mit dem SAS Software

Navigator (s. Abb. 2.1) leicht zu installieren. Die Installation des SAS-ODBC-Treibers soll zunächst, ohne SAS-Datenquellen zu definieren, abgeschlossen werden.

Der aktuelle SAS-ODBC-Treiber 9.1 kann auch von der SAS-Website http://www.sas.com/apps/demosdownloads/odbdriver_PROD_9.1_sysdep.jsp;jsessionid=34F76C8F27CEFD8970AF3EFA60FDB8DA.tomcat3?packageID=000300 heruntergeladen werden.

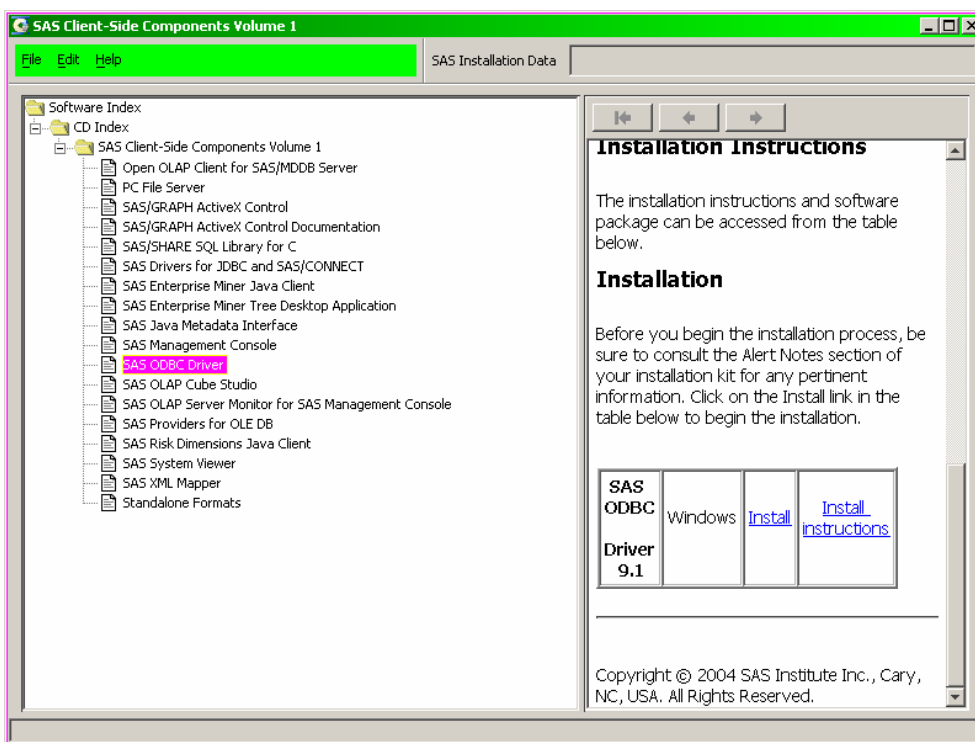


Abb. 2.1: SAS Client-Side Components Volume 1

ODBC-Datenquellen werden über den ODBC-Datenquellen-Administrator definiert, der in der Regel bei der Installation von MS Office oder der Windows-Systemkomponente MDAC (Microsoft Data Access Components) mit installiert wird. Der ODBC-Datenquellen-Administrator ist bei deutschsprachigen Betriebssystemen in der Regel auch deutschsprachig. Ist bei der Installation von SAS allerdings die Windows-Systemkomponente MDAC nicht in der von SAS 9 geforderten Aktualität verfügbar, dann muss vor der eigentlichen SAS-Installation diese Komponente nachinstalliert werden. Geschieht das von der SAS-Installations-CD, dann ist der entsprechende ODBC-Datenquellen-Administrator englischsprachig.

Zur Demonstration der Definition einer SAS-ODBC-Datenquelle wird folgendes angenommen:

- SAS 9.1.3 ist in dem Ordner installiert, der bei der Installation als Voreinstellung angeboten wird (C:\Programme\SAS\SAS 9.1)
- In SAS ist die permanent verfügbare library **saslib** definiert, die den Ordner d:\daten\saslib repräsentiert, der die SAS-Datendatei **dat.sas7bdat** enthält.

Zur Definition einer SAS-ODBC-Datenquelle mit dem Namen **sasdq**, über die dann auf die SAS-Datei **dat** zugegriffen werden kann, sind folgende Schritte notwendig:

1. Öffnen des ODBC-Datenquellen-Administrators (Windows XP Professional: **Start** → **Einstellungen** → **Systemsteuerung** → **Verwaltung** → **Datenquellen (ODBC)** → **ODBC-Datenquellen-Administrator** (s. Abb. 2.2)

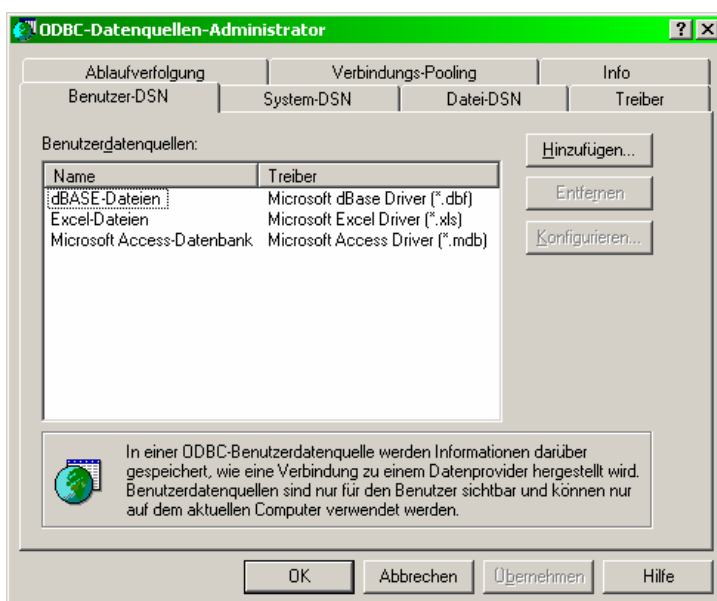


Abb. 2.2: ODBC-Datenquellen-Administrator

2. **ODBC-Datenquellen-Administrator: Hinzufügen** → **Neue Datenquelle erstellen** (s. Abb. 2.3)

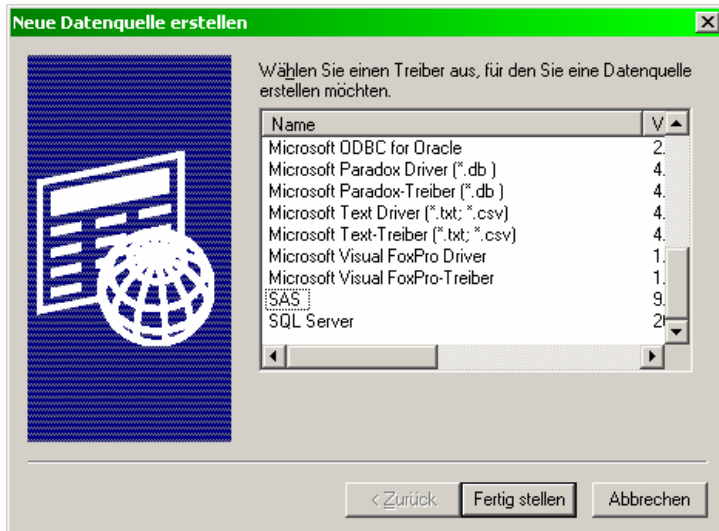


Abb. 2.3: Neue Datenquelle erstellen

3. **Neue Datenquelle erstellen: SAS** → **Fertig stellen** → **SAS ODBC Driver Configuration** (s. Abb. 2.4). In diesem Fenster sind die Schaltflächen General, Servers und Libraries vorhanden. Da mindestens ein Server definiert werden muss, wird zunächst die Schaltfläche *Servers* gewählt (s. Abb. 2.5).

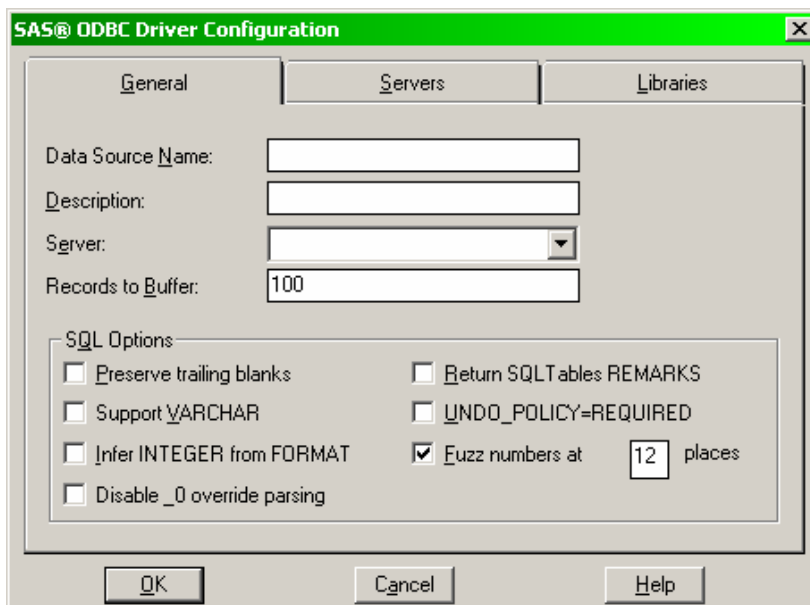


Abb. 2.4: SAS ODBC Driver Configuration (General)

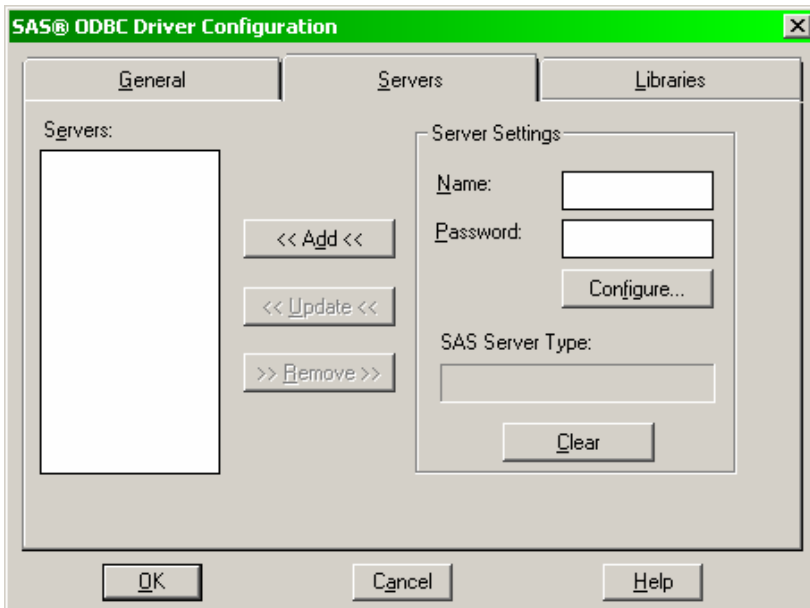


Abb. 2.5: SAS ODBC Driver Configuration (Servers)

4. Durch den Eintrag (des frei wählbaren, maximal 8 Zeichen langen) Servernamens **sasserv** im Feld Name und anschließende Wahl von **Configure** gelangt man zum Fenster **Local Options** mit den Voreinstellungen in Abb. 2.6. SAS 9.1.3 erkennt, dass es sich um ein deutschsprachiges Betriebssystem handelt und wählt automatisch die richtigen Pfadangaben, wenn bei der SAS-Installation auch die Standardvorgaben benutzt wurden. Mit **OK** wird die Einstellung der Local Options beendet und durch **Add** die Serverdefinition (s. Abb. 2.7) abgeschlossen.

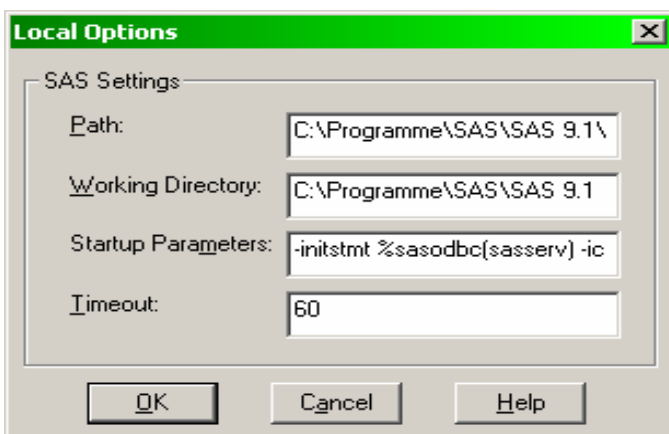


Abb. 2.6: Local Options (Voreinstellungen)

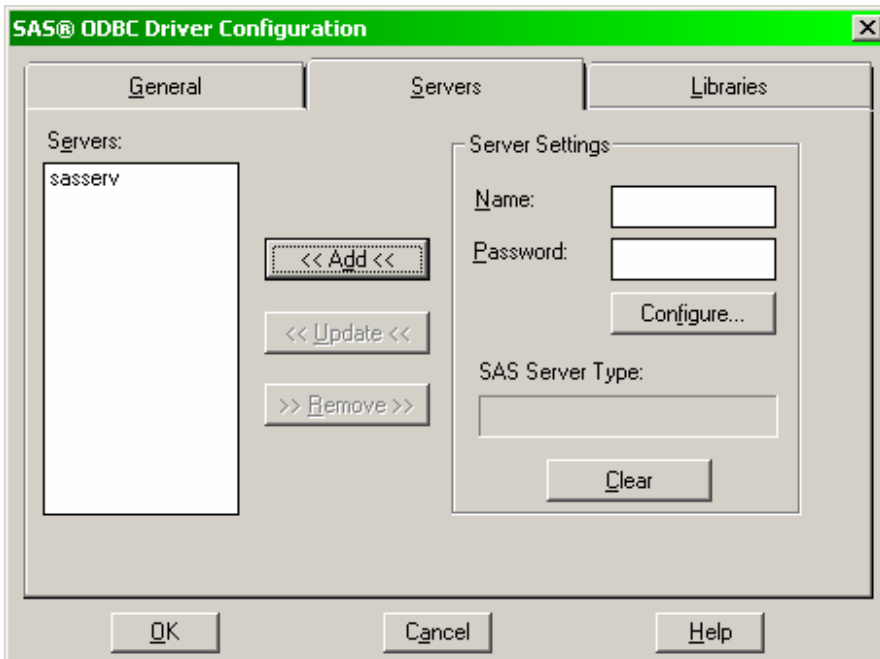


Abb. 2.7: SAS ODBC Driver Configuration (Servers: sasserv)

5. Im Fenster (Libraries) sind jetzt der Name der library (**saslib**) und unter Host File die zur library **saslib** gehörige Pfadangabe (**d:\daten\saslib**) einzutragen. Diese Eintragungen werden mit **Add** abgeschlossen (s. Abb. 2.8).

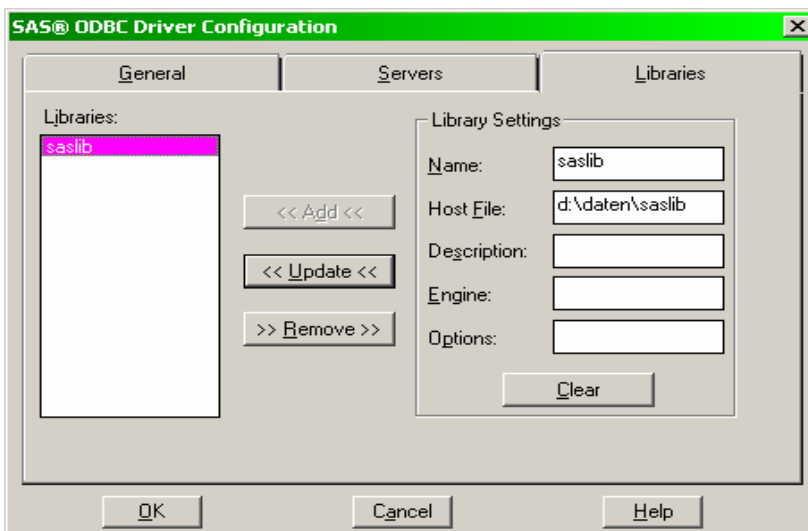


Abb. 2.8: SAS ODBC Driver Configuration (Libraries)

- Nach Eintragung des (frei wählbaren) Datenquellennamens **sasdq** im Fenster **SAS ODBC Driver Configuration** (General), vgl. Abb. 2.9, wird mit OK die Definition der SAS-ODBC-Datenquelle abgeschlossen. Im ODBC-Datenquellen-Administrator erscheint die definierte SAS-ODBC-Datenquelle in der Liste der Benutzer-DSN (s. Abb. 2.10). Mit **OK** kann der ODBC-Datenquellen-Administrator verlassen werden.

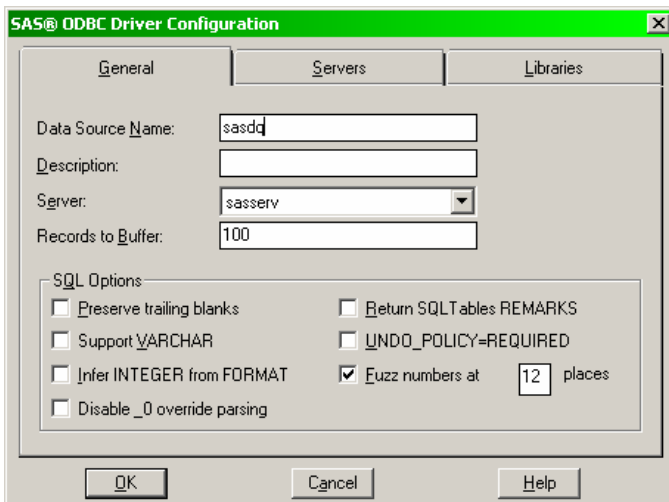


Abb. 2.9: SAS ODBC Driver Configuration

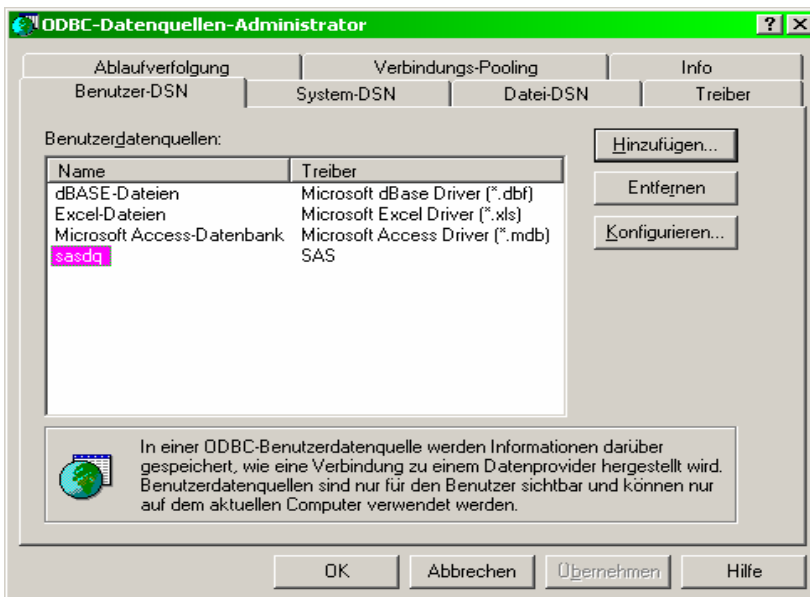


Abb. 2.10: ODBC-Datenquellen-Administrator (sasdq)

7. Der definierte SAS-Server **sasserv** muss abschließend noch mit einem Eintrag in die Windows-Systemdatei **Services** (Windows XP: C:\WINDOWS\system32\drivers\etc) auf einen freien Port oberhalb 1024 und das Protokoll **tcp** (Transmission Control Protocol) aktiviert werden (s. Abb. 2.11). Dieser Eintrag kann mit einem beliebigen Texteditor erfolgen.

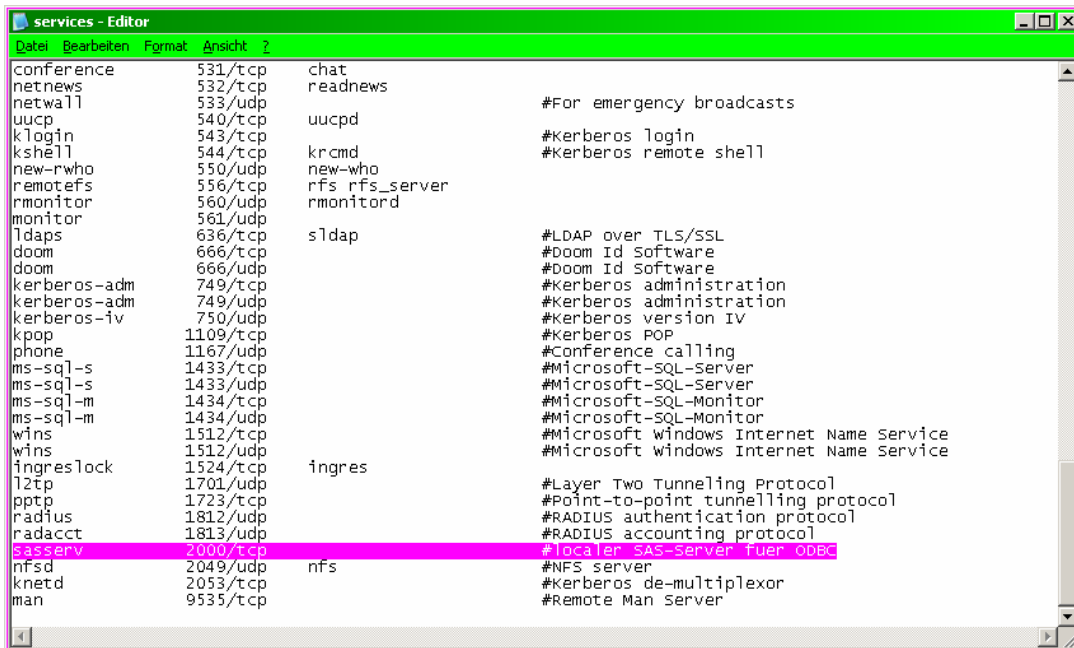


Abb. 2.11: Systemdatei Services

Die Definition von Excel- und Access-Datenquellen mit dem ODBC-Datenquellen-Administrator ist weitgehend selbsterklärend und wird deshalb hier nicht weiter beschrieben.

3 Import von SAS-Daten in MS Excel / Access 2003

Für die in diesem Kapitel beschriebenen Datenimporte soll die in Kapitel 2 definierte SAS-ODBC-Datenquelle **sasdq** und die SAS-Datei **dat.sas7bdat** verwendet werden. Der Inhalt dieser Datei (4 Variablen: zwei vom Typ Character und jeweils eine vom Typ Datum bzw. numerisch) ist in der Abb. 3.1 dargestellt.

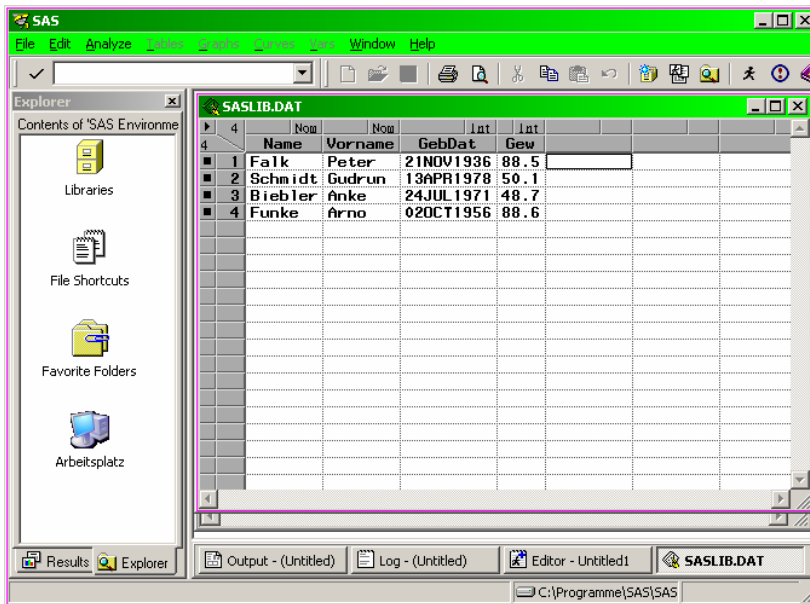


Abb. 3.1: Beispieldaten in der SAS-Datei **dat**

3.1 Import von SAS-Daten in Excel

Der Import der SAS-Datei `saslib.dat` in MS Excel 2003 mittels ODBC wird mit den folgenden Schritten realisiert:

1. Excel starten. (Es wird eine neue Excel-Datei erstellt mit dem Namen Mappel1.) **Daten** → **Externe Daten importieren** → **Daten importieren** → **Datenquelle auswählen** (s. Abb. 3.2)

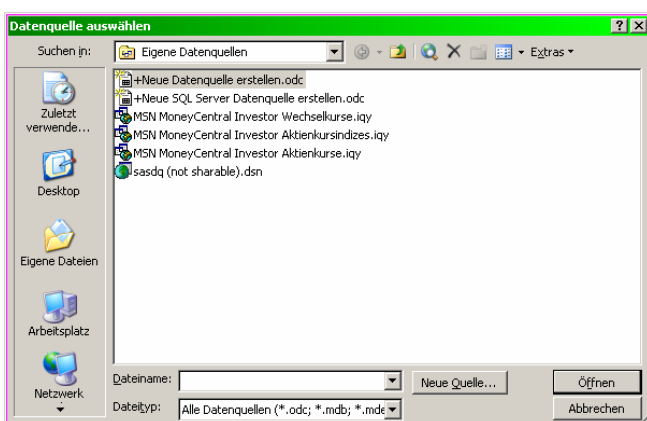


Abb. 3.2: Datenquelle auswählen

2. **Datenquelle auswählen**: *sasdq* → **Öffnen** → (SAS startet) **Tabelle auswählen** (s. Abb. 3.3) (Dieses Fenster zeigt alle verfügbaren SAS-Tabellen.)

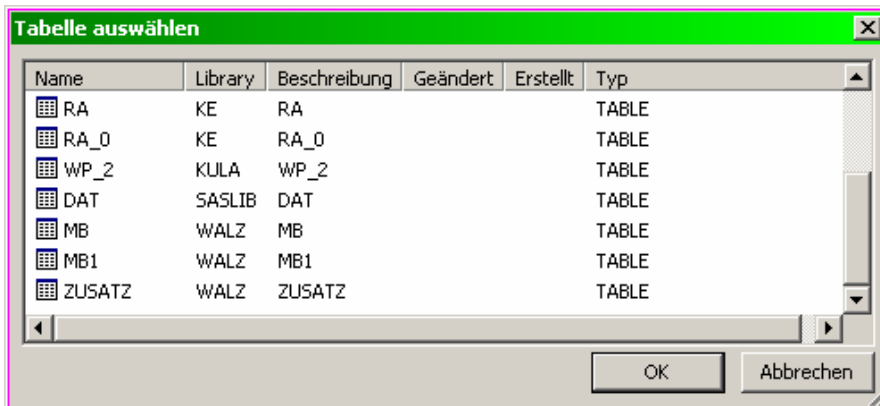


Abb. 3.3: Tabelle auswählen

3. **Tabelle auswählen**: *DAT* → **Daten importieren** (s. Abb. 3.4).

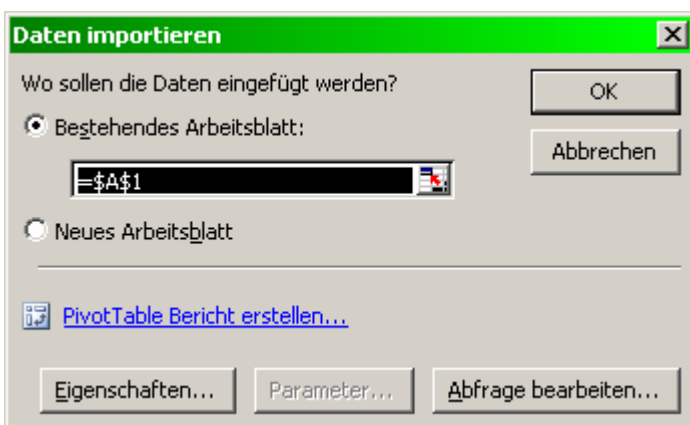


Abb. 3.4: Daten importieren

4. Über das Fenster **Daten importieren** kann man sich über die Schaltfläche ***Eigenschaften*** die Eigenschaften des externen Datenbereiches anzeigen lassen oder auch über ***Abfrage bearbeiten*** die im Hintergrund mit MS Query erstellte SQL-Abfrage, z. B. zur speziellen Auswahl von Daten aus der SAS-Tabelle ***dat***, verändern. Mit ***OK*** werden die Daten aus der SAS-Datei ***dat*** ab Spalte A und Zeile 1 in Tabelle1 der geöffneten Excel-Datei übernommen. Neben den in der SAS-Datei verfügbaren Variablen wird auch die SAS-System-Variable ***_OBSTAT_*** übernommen. Diese Variable ist in Excel leicht zu löschen. Eine

Übernahme kann auch durch vorheriges Löschen über *Abfrage bearbeiten* verhindert werden.

5. Das Speichern der Excel-Datei und das Beenden von SAS schließen den Datentransfer ab.

Die erzeugte Excel-Datei soll hier mit dem Namen **Dat.xls** unter **d:\daten\excel** abgespeichert werden, um sie für den im Abschnitt 4.1 beschriebenen Datentransfer zu verwenden.

3.2 Import von SAS-Daten in Access

Zur Realisierung des Datentransfers der SAS-Daten aus der Datei **saslib.dat** in eine Tabelle der (leeren) Access-Datenbank **dat.mdb** mittels ODBC und der SAS-ODBC-Datenquelle **sasdq** sind die folgenden Schritte notwendig:

1. Starten von MS Access und bestehende Datenbank **dat.mdb** unter **d:\daten\access** öffnen → **dat: Datenbank**
2. **dat : Datenbank: Tabellen** → *Neu* → **Neue Tabelle**
3. **Neue Tabelle: Tabelle importieren** → *OK* → **Importieren**
4. **Importieren**: Dateityp: *ODBC-Datenbanken* → **Datenquelle auswählen**
5. **Datenquelle auswählen**: *Computerdatenquelle* → *sasdq* → *OK* → (SAS startet) **Objekte importieren**
6. **Objekte importieren**: *SASLIB.DAT* → *OK* → **SASLIB_DAT** (als neue Tabelle in der Access-Datenbank **dat.mdb**)
7. Beenden von Access und SAS

4 Import von Excel- / Access-Daten in SAS

4.1 Import von Excel-Daten in SAS

Zur Demonstration des Imports von Excel-Daten in SAS werden die Daten der Tabelle1 der im Abschnitt 3.1 erzeugten Exceldatei Dat.xls verwendet. Um den Datenimport mit dem folgenden SAS-Programm durchführen zu können, muss zunächst mit dem Datenquellenadministrator eine mit der Exceldatei Dat.xls verknüpfte Excel-Datenquelle definiert werden. Der Name dieser Datenquelle sei **exceldq**. Der eigentliche Datenimport erfolgt dann mit der SAS-Prozedur SQL in einem entsprechenden SAS-Programm.

SAS-Programm zum Import von Excel-Daten in SAS

```
/* 1. ODBC-Verbindung mit der Excel-Datenquelle exceldq über proc sql */  
  
proc sql;  
    connect to odbc (dsn=exceldq);  
  
/* 2. Erstellen der SAS-Datei dat1 und einfügen der Daten aus Tabelle1 */  
  
    create table dat1 as select * from connection to odbc  
    (Excel-Abfrage);  
  
/* 3. Schließen der ODBC-Verbindung und beenden von proc sql */  
  
    disconnect from odbc;  
quit;
```

Um das angegebene Programm lauffähig zu machen, ist der Ausdruck **Excel-Abfrage** durch den tatsächlichen SQL-Query-Code zu ersetzen. Dieser SQL-Code (Excel-Query) ist relativ leicht erhältlich. Dazu sind in Anlehnung an die im Abschnitt 3.1 beschriebene Vorgehensweise beim Import von SAS-Daten in Excel folgende Schritte notwendig:

1. Excel starten (Es wird eine neue Excel-Datei mit dem Namen Mappel geöffnet.) → **Daten** → **Externe Daten importieren** → **Neue Abfrage erstellen** → (MS Query startet mit dem Fenster:) **Datenquelle auswählen**
2. **Datenquelle auswählen**: **Datenbanken** → **Excel-Dateien** → **OK** → **Arbeitsmappe auswählen**
3. **Arbeitsmappe auswählen**: Datenbankname: **Dat.xls** → **OK** → **Query-Assistent – Spalten auswählen** (ein Fenster, das die in der Datei Dat.xls, Tabelle1 vorhandenen Variablen (Spalten) anzeigt, s. Abb. 4.1)

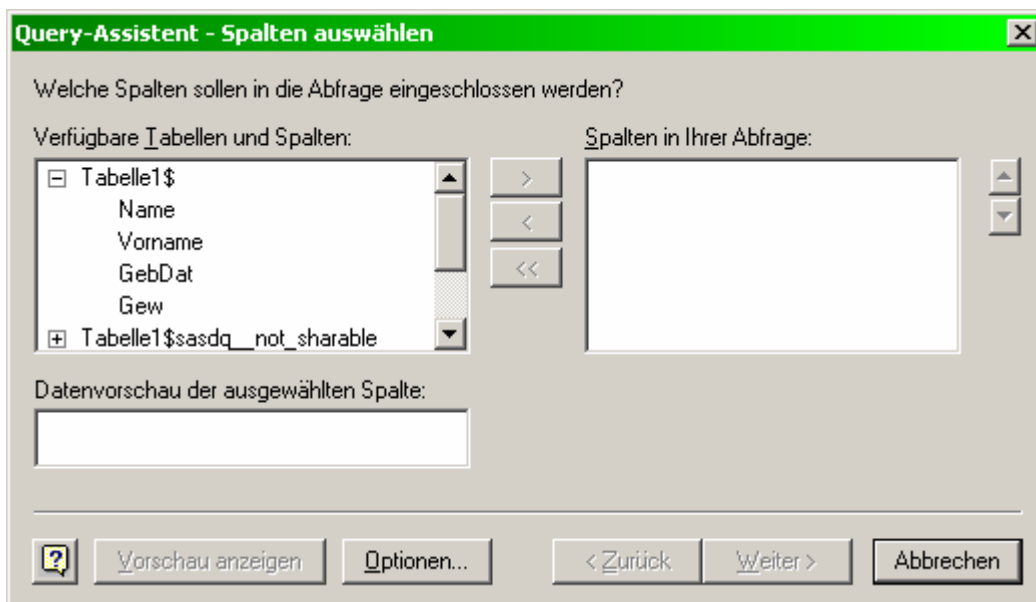


Abb. 4.1: Query-Assistent – Spalten auswählen

4. **Query-Assistent – Spalten auswählen**: Übernahme der gesamten Tabelle1 oder einzelner Spalten: **Weiter** → **Query-Assistent Daten filtern**: **Weiter** → **Query-Assistent Sortierreihenfolge**: **Weiter** → **Query-Assistent Fertig stellen**: **Daten in MS Query ansehen oder bearbeiten** → **Fertig stellen** → **Microsoft Query** (s. Abb. 4.2)

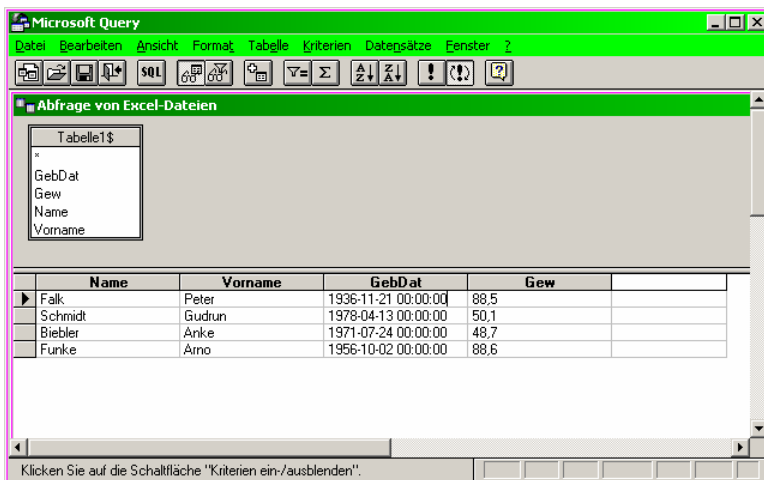


Abb. 4.2: Microsoft Query

5. Microsoft Query: **SQL** → **SQL** (enthält den SQL-Code der Abfrage, s. Abb. 4.3)

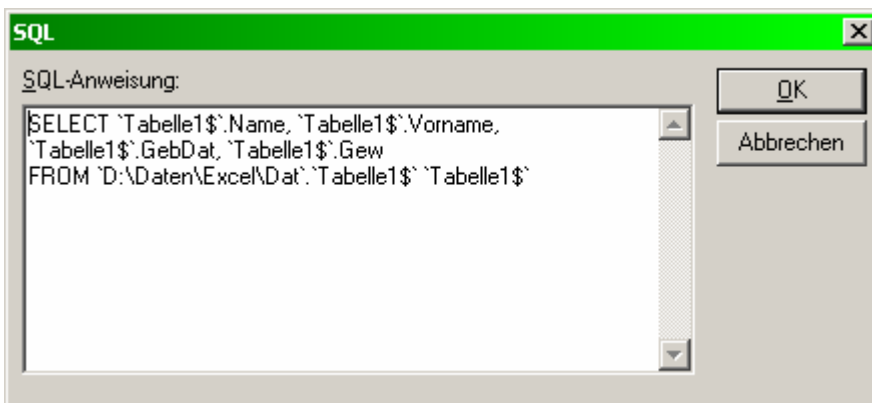


Abb. 4.3: SQL-Fenster

6. SQL: Kopieren des SQL-Codes in die Zwischenablage → **Abbrechen** → Microsoft Query → **Schließen** (×) → **Abbrechen** → Excel beenden (ohne Speichern)
7. **SQL-Code** aus der Zwischenablage an die Stelle von **Excel-Abfrage** in obiges SAS-Programm kopieren

Nach Abarbeitung dieser 7 Schritte ist das SAS-Programm lauffähig und realisiert den gewünschten Datenimport.

Allerdings wird beim Import der Daten in SAS bei den Character-Variablen automatisch eine Breite von 200 Zeichen gesetzt. Dieser Fehler bedarf noch einer Aufklärung.

4.2 Import von Access-Daten in SAS

Der Import von Access-Daten in SAS soll am Beispiel der im Abschnitt 3.2 erzeugten Access-Tabelle `SASLIB_DAT` demonstriert werden. Um den Datenimport mit dem unten angegebenen SAS-Programm durchführen zu können, muss zunächst mit dem Datenquellenadministrator eine Access-Datenquelle definiert werden. Diese Datenquelle soll **accessdq** heißen und muss mit der Access-Datenbank `dat.mdb`, in der sich die zu importierende Tabelle `SASLIB_DAT` befindet, verknüpft sein.

Die Abarbeitung des folgenden SAS-Programms realisiert den Datenimport, ohne dass MS Access gestartet wird.

SAS-Programm zum Import von Access-Daten in SAS:

```
/* 1. ODBC-Verbindung mit der Access-Datenquelle accessdq über proc sql */  
  
proc sql;  
    connect to odbc (dsn=accessdq);  
  
/* 2. Erstellen der (temporären) SAS-Datei dat2 und Einfügen der Daten aus der  
   Access-Tabelle SASLIB_DAT */  
  
    create table dat2 as select * from connection to odbc  
    (select * from [saslib_dat]);  
  
/* 3. Schließen der ODBC-Verbindung und Beenden von proc sql */  
  
    disconnect from odbc;  
quit;
```


Literatur

- [1] Rudolph, P. E. (2000): ODBC zwischen SAS 6.12 und MS Access / Excel 2000. SAS-Anwenderhandbuch im Netz.
<http://www.urz.uni-heidelberg.de/statistik/sas-ah/>
- [2] Rudolph, P. E.; Geissler, H.; Ortseifen, C. (1998): Datenaustausch zwischen SAS und anderen Anwendungen mit DDE und ODBC. Konferenzbeiträge der 2. Konferenz für SAS-Benutzer in Forschung und Entwicklung (KSFE), Friedrich-Schiller-Universität Jena, S. 200-220.
- [3] Rudolph, P. E.; Ortseifen, C.; Havemann, P. (2001): Datenaustausch zwischen SAS 8 und anderen Windows-Applikationen mittels ODBC. Proceedings der 5. Konferenz für SAS-Benutzer in Forschung und Entwicklung (KSFE), Universität Hohenheim, S. 289-298
- [4] SAS Institute Inc. (2004): SAS/ACCESS 9.1 Supplement for ODBC (SAS/ACCESS for Relational Databases). Cary, NC: SAS Institute Inc.
- [5] SAS Institute Inc. (2004): SAS ODBC Driver 9.1: User's Guide and Programmer's Reference. Cary, NC: SAS Institute Inc.