

Datentransfer zwischen SAS 8.2 und MS SQL-Server bzw. Oracle mit ODBC

Paul Eberhard Rudolph,

Peter Havemann

Forschungsinstitut für die
Biologie landw. Nutztiere

18196 Dummerstorf

rudolph@fbn-dummerstorf.de

havemann@fbn-dummerstorf.de

Michael Wodny

E.-M.-Arndt-Univ. Greifswald

Inst. f. Biometrie u. Med. Inf.

17487 Greifswald

wodny@biometrie.uni-greifswald.de

Zusammenfassung

Datenbanksysteme werden auch in der Forschung immer mehr zum Halten und Pflegen von Daten verwendet. Statistische Auswertungen derartiger Daten mit SAS setzen voraus, dass es gut handhabbare Methoden gibt, um die Datenbankdaten in SAS-spezifische Datendateien einzulesen und Ergebnisse auch wieder in die Datenbank zurückzuschreiben. SAS bietet dazu mit den „SAS/ACCESS Interface to ...“-Produkten datenbankspezifische Schnittstellen für eine ganze Reihe von Datenbanksystemen an.

Eine weitere derartige Schnittstelle ist die von Microsoft entwickelte Open Data Base Connectivity (ODBC). ODBC hat den Vorteil, dass mit der Lizenzierung von SAS/ACCESS to ODBC und der Verfügbarkeit eines ODBC-Treibers für die entsprechende Datenbank der Zugriff auf Datenbankdaten möglich wird, ohne die datenbankspezifischen SAS/ACCESS-Produkte zu lizenzieren. Eine ausführliche Einführung in diese Technik ist außer in den entsprechenden SAS-Dokumentationen auch bei RUDOLPH u. a. [3] (1998), RUDOLPH u. a. [5] (2001), hier hauptsächlich auf den Datenaustausch zwischen SAS und MS Excel bzw. MS Access orientiert, zu finden.

Am Beispiel der Datenbanksysteme Microsoft SQL Server und Oracle wird gezeigt, wie der Datenaustausch zwischen SAS und Datenbanksystemen mittels ODBC funktioniert.

Keywords: Datenaustausch, SAS, MS SQL Server, Oracle, ODBC .

1 Einleitung

Open Database Connectivity (ODBC) ist eine von Microsoft entwickelte Programmierschnittstelle, um Daten zwischen verschiedenen Anwendungen austauschen zu können. Insbesondere können Anwendungen (z.B. SAS) mit ODBC auf Daten in Datenbankmanagementsystemen zugreifen, die SQL (Structured Query Language) als Standardsprache für den Datenzugriff unterstützen.

In der online-Hilfe des Betriebssystems Windows 2000 Professional (deutsch) wird ODBC merkwürdigerweise mit Leistungsindikatoren übersetzt.

Um einen Datenaustausch mittels ODBC durchzuführen, müssen außer ODBC-fähigen Anwendungen (z.B. SAS mit der Komponente ACCESS TO ODBC, MS SQL Server, Oracle) die ODBC-Treiber der Anwendungen verfügbar sein. Über den ODBC-Datenquellen-Administrator (Windows-Systemsteuerung) können dann ODBC-Datenquellen (Data Sources) definiert werden, über die der eigentliche Zugriff auf die Daten erfolgt.

Eine ausführliche Einführung in diese Technik findet man außer in den Dokumentationen der entsprechenden Anwendungen auch bei Rudolph u.a. [3] (1998). Um aus SAS auf Daten in einer MS SQL-Server-Datenbank bzw. Oracle-Datenbank zuzugreifen und auch Daten aus SAS wieder in die Datenbank zurückzuschreiben, ist neben der Installation der SAS-Komponente ACCESS TO ODBC nur der ODBC-Treiber für MS SQL Server bzw. Oracle erforderlich. Der SAS-ODBC-Treiber wird nur benötigt, wenn aus einer anderen Anwendung auf SAS-Datendateien zugegriffen werden soll.

In Kapitel 2 wird an einem Beispiel ausführlich beschrieben, wie man aus SAS mittels ODBC auf Daten in einer MS SQL-Server-Datenbank zugreifen und auch Daten wieder in die SQL-Server-Datenbank zurückschreiben kann. Der Zugriff aus SAS auf Oracle-Datenbanken erfolgt im Prinzip analog zu dem auf MS SQL-Server-Datenbanken. Deshalb wird in Kapitel 3 diese Vorgehensweise auch nur kurz beschrieben.

Voraussetzung für das Funktionieren der Datentransfers ist eine Windows-Version von SAS 8.2 mit den Komponenten SAS/CORE, SAS/BASE und SAS/ACCESS to ODBC.

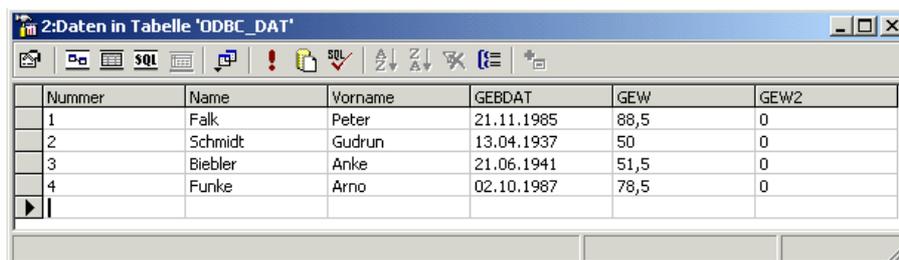
Die dargestellten Bildschirm- bzw. Fensterinhalte wurden unter Windows 2000 Professional (deutsch) und MS SQL Server Version 7.0 (deutsch) bzw. Windows 98 und Oracle 8i erstellt. Unter Windows 95/98/Me/NT/XP können die entsprechenden Inhalte eventuell von den hier dargestellten leicht abweichen. Bei den beschriebenen Vorgehensweisen sind Fenstertitel durch Unter-

streichen/fett und die Auswahl aus in diesen Fenstern vorhandenen Schaltflächen oder Menüs durch Kursivschrift/fett gekennzeichnet.

2 Datentransfer zwischen SAS und MS SQL-Server

Zur Beschreibung des Datentransfers zwischen SAS und MS SQL Server soll folgendes angenommen werden:

- Der PC, auf dem SAS installiert ist, sei in einem lokalen Netz mit einem Server mit dem Namen **nt0102** verbunden.
- Auf dem Server liegt die MS SQL-Server-Datenbank **pedat** mit der Tabelle **odbc_dat** (s. Abb. 2.1), für die der Nutzer des SAS-PC Lese- und Schreibzugriff über die standardmäßigen Anmeldedaten (NT-Authentifizierung) für das lokale Netz besitzt.



Nummer	Name	Vorname	GEBDAT	GEW	GEW2
1	Falk	Peter	21.11.1985	88,5	0
2	Schmidt	Gudrun	13.04.1937	50	0
3	Biebler	Anke	21.06.1941	51,5	0
4	Funke	Arno	02.10.1987	78,5	0

Abbildung 2.1: MS SQL-Server-Datenbank pedat, Tabelle odbc_dat

In Rudolph u.a. [5] (2001) wurde bereits beschrieben, wie man mit der SAS-Prozedur SQL Daten aus einer Tabelle in einer SQL-Server-Datenbank einlesen und verändern kann. In Kapitel 2.1 wird das entsprechende SAS-Programm noch einmal dargestellt. Die im weiteren beschriebenen Möglichkeiten des Datentransfers erfordern zunächst die Definition einer ODBC-Datenquelle mit dem ODBC-Treiber für den MS SQL Server. Diese Datenquellendefinition wird in Kapitel 2.2 erläutert. In Kapitel 2.3 wird der Datenaustausch über eine interaktive Bibliotheksdefinition beschrieben. Kapitel 2.4 schließlich befasst sich kurz mit dem Datenzugriff über das SAS SQL Query Window.

2.1 Datentransfer mit PROC SQL

Als Beispiel sollen folgende Aufgaben gelöst werden:

- Erzeugen der temporären SAS-Datei data aus den Daten der Tabelle odbc_dat in der MS-SQL-Server-Datenbank pedat
- Ersetzen der in der Spalte GEW2 in der Tabelle odbc_dat vorhandenen Werte durch die Werte 2*GEW

Diese Aufgaben kann das folgende SAS-Programm lösen:

```
/* 1. (ODBC-)Verbindung mit der SQL-Server-
   Datenbank pedat auf dem Rechner nt0102
   (im Intranet) über proc sql */

proc sql;
  connect to sqlservr (server = nt0102
                     database = pedat
                     language = deutsch) ;

/* 2. Erstellen der (temporären) SAS-Datei data
   und Einfügen der Daten aus der Tabelle
   odbc_dat */

  create table data as select * from connection
  to sqlservr (select * from [odbc_dat]);

/* 3. Aktualisieren der in der Tabelle odbc_dat
   vorhandenen Spalte gew2 der SQL-Server-
   Datenbank pedat */

  execute (update [odbc_dat]
  set gew2=2*gew) by sqlservr;

/* 4. Schließen der(ODBC-)Verbindung und beenden
   von procsql */

  disconnect from sqlservr;
quit;
```

Bemerkung 2.1:

Für den Datenaustausch mit dem obigen SAS-Programm ist es nicht notwendig, eine ODBC-Datenquelle mit dem ODBC-Treiber für den MS SQL Server zu definieren.

Im connect-Statement des obigen SAS-Programms war die (sonst übliche Angabe) von Benutzername und Passwort nicht notwendig, da der Zugang vom

SAS-PC zum Server nt0102 und der entsprechenden SQL-Datenbank über die NT-Authentifizierung im lokalen Netz erfolgte. Das execute-Statement des SAS-Programms muss mit dem SQL-Sprachumfang von MS SQL Server verträglich sein.

2.2 ODBC-Datenquellendefinition für MS SQL Server

Zur Definition einer ODBC-Datenquelle wird (in Windows 2000) über die Wahl von

Start → **Einstellungen** → **Systemsteuerung** → **Verwaltung** → **Datenquellen (ODBC)**

der ODBC-Datenquellen-Administrator aufgerufen (s. Abb. 2.2).

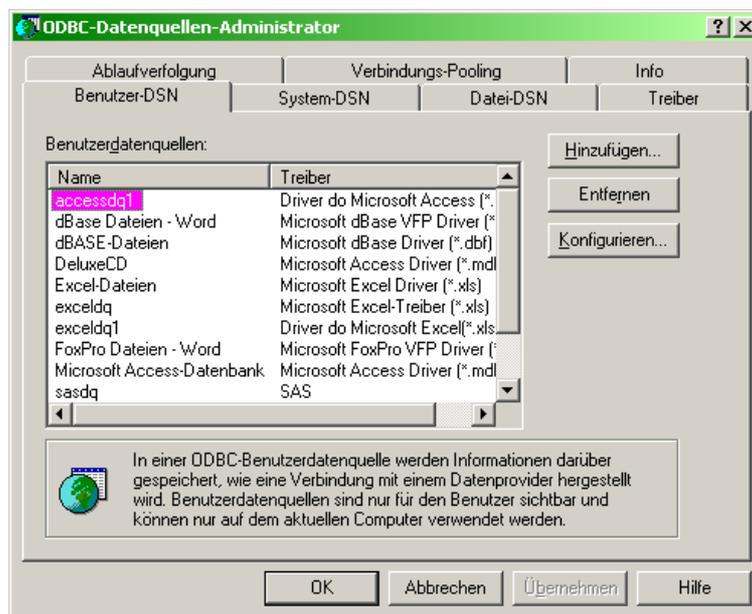


Abbildung 2.2: ODBC-Datenquellen-Administrator

Im Modus Benutzer- oder System-DSN (**Data Source Name**) gelangt man über die Schaltfläche **Hinzufügen** zu einem Fenster mit dem Titel Neue Datenquelle erstellen (s. Abb. 2.3).

Die Wahl des ODBC-Treibers für SQL Server und **Fertig stellen** führt zum Fenster Neue Datenquelle für SQL Server erstellen (s. Abb. 2.4).

In diesem Fenster werden ein (beliebig wählbarer) Name (hier sql_dq) für die Datenquelle und der entsprechende Server (hier NT0102 aus der verfügbaren Liste gewählt) eingetragen. Mit **Fertigstellen** ist die Definition der ODBC-SQL-Server-Datenquelle abgeschlossen.



Abbildung 2.3: Neue Datenquelle erstellen

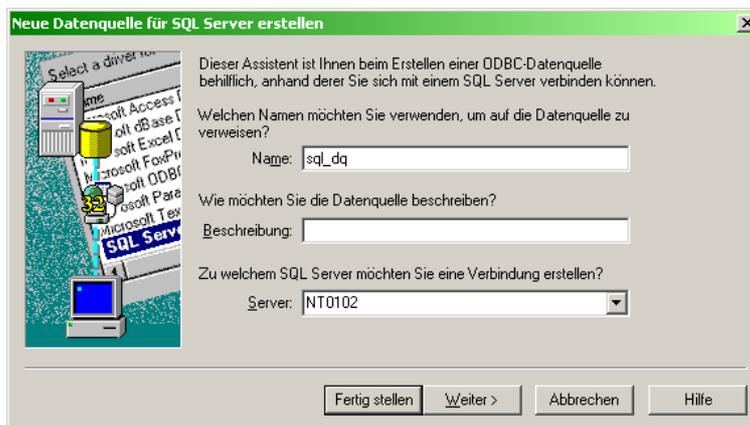


Abbildung 2.4: Neue Datenquelle für SQL Server erstellen

2.3 Interaktive Bibliotheksdefinition

Nach der Definition einer ODBC-SQL-Server-Datenquelle kann auf folgende Weise interaktiv eine Bibliothek (library) definiert werden. Durch Anklicken der New-Library-Ikone in SAS öffnet sich das Fenster New Library (s. Abb. 2.5).

In diesem Fenster ist ein Name für die neue Bibliothek einzutragen und für Engine ODBC aus der verfügbaren Liste zu wählen. In der Zeile Data Source wird die verfügbare ODBC-SQL-Server-Datenquelle mit dem Namen **sql_dq** ausgewählt. Auf die Angaben Benutzer (User ID) und Passwort kann aus den in Bemerkung 2.1 genannten Gründen verzichtet werden. Soll die so definierte

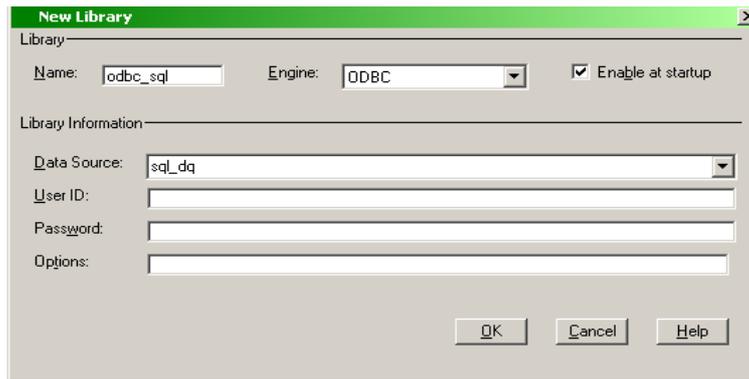


Abbildung 2.5: New Library

Bibliothek auch in folgenden SAS-Sitzungen verfügbar sein, ist das Kästchen **Enable at startup** zu aktivieren. Mit **OK** wird das Fenster verlassen und die Bibliotheksdefinition ist abgeschlossen.

Die Tabelle **odbc_dat** in der SQL-Server-Datenbank **pedat** kann in SAS jetzt wie eine normale SAS-Datendatei verwendet werden unter einem zweistufigen Namen. Dieser Name setzt sich aus dem Bibliotheksnamen (odbc_sql) und dem Tabellennamen (odbc_dat) zusammen, also hier: **odbc_sql.odbc_dat**. In der Abbildung 2.6 wird das mittels SAS/INSIGHT illustriert.

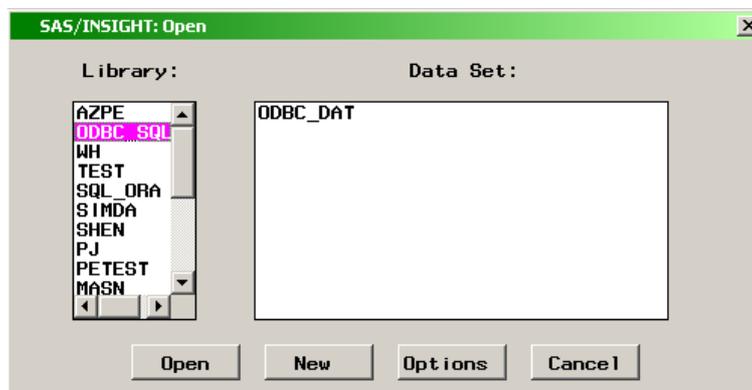


Abbildung 2.6: SAS/INSIGHT: Open

Bemerkung 2.2:

Die interaktive Bibliotheksdefinition kann jederzeit durch ein einem SAS-Data- oder Proc-Step vorangehendes SAS/ACCESS LIBNAME Statement ersetzt werden. Um z.B. die temporäre SAS-Datei **dat** mit den Daten der Tabelle **odbc_dat** aus der MS SQL-Server-Datenbank **pedat** zu erstellen, kann folgender SAS-Code abgearbeitet werden:

```
libname odbc_sql odbc dsn=sql_dq /* uid=   pwd= */ ;  
data dat;  
    set odbc_sql.odbc_dat;  
run;
```

2.4 Datentransfer mittels SAS SQL Query Window

Eine weitere Möglichkeit, aus SAS auf Daten in einer SQL-Server-Datenbank zuzugreifen, besteht in der Nutzung des SAS SQL Query Windows. Das SAS SQL Query Window ist eine interaktive Schnittstelle, die auch einem nicht mit SQL vertrauten Anwender gestattet, Datenbankabfragen zu generieren, zu speichern und abzuarbeiten. Bei der Nutzung des SQL Query Windows wird im Hintergrund der entsprechende Programmcode für die SAS-Prozedur SQL erzeugt.

Wählt man in SAS *Tools* → *Query*, dann öffnet sich das Fenster mit dem Titel **SQL Query Tables** (s. Abb. 2.7). In diesem Fenster sind zunächst alle in SAS definierten Bibliotheken aufgelistet.

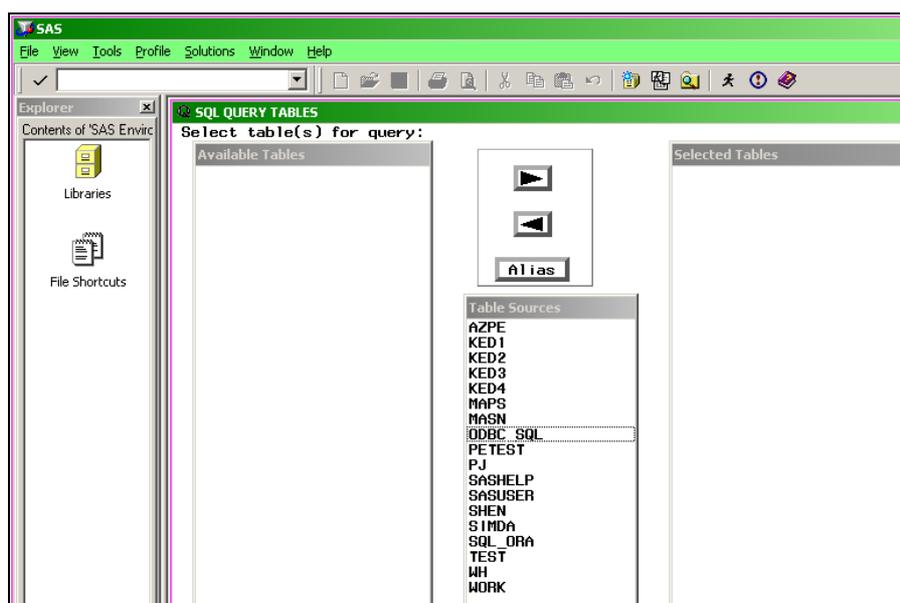


Abbildung 2.7: SQL Query Tables

Hat man bereits interaktiv oder über ein Libname Statement eine ODBC-Bibliothek (mit der entsprechenden SQL-Datenquelle) definiert, dann erscheint diese (ODBC_SQL) natürlich auch hier. Bei Auswahl dieser Bibliothek erscheinen in der Spalte „tabel(s) for query“ die dort verfügbaren Tabellen. Nach dem Markieren einer Tabelle (z.B. ODBC_DAT) in dieser Spalte kann über die entsprechende Pfeil-Schaltfläche diese Tabelle in die Spalte „Selected Tables“

angezeigt werden. Ein Doppelklick auf die gewählte Tabelle liefert das gleiche Resultat. Wählt man **Columns** bei markierter Tabelle in der Spalte „Selected Tables“ im über die rechte Maustaste erreichbaren Menü, dann öffnet sich ein neues Fenster mit dem Titel **SQL Query Columns** (s. Abb. 2.8). In der linken Spalte „Select Columns for Query“ sind die in der Tabelle `odbc_dat` vorhandenen Variablen aufgelistet. Die für den Datentransfer gewünschten Variablen (z.B. Nummer, Name Gew, (s. Abb. 2.8) können jetzt über die Pfeilschaltflächen oder Doppelklick ausgewählt werden. Nach der Wahl von **File** → **Create Table from Query Results** können in einem weiteren Fenster ein Bibliotheksname und ein Dateiname angegeben bzw ausgewählt werden. Mit **OK** in diesem Fenster wird der Datentransfer abgeschlossen.

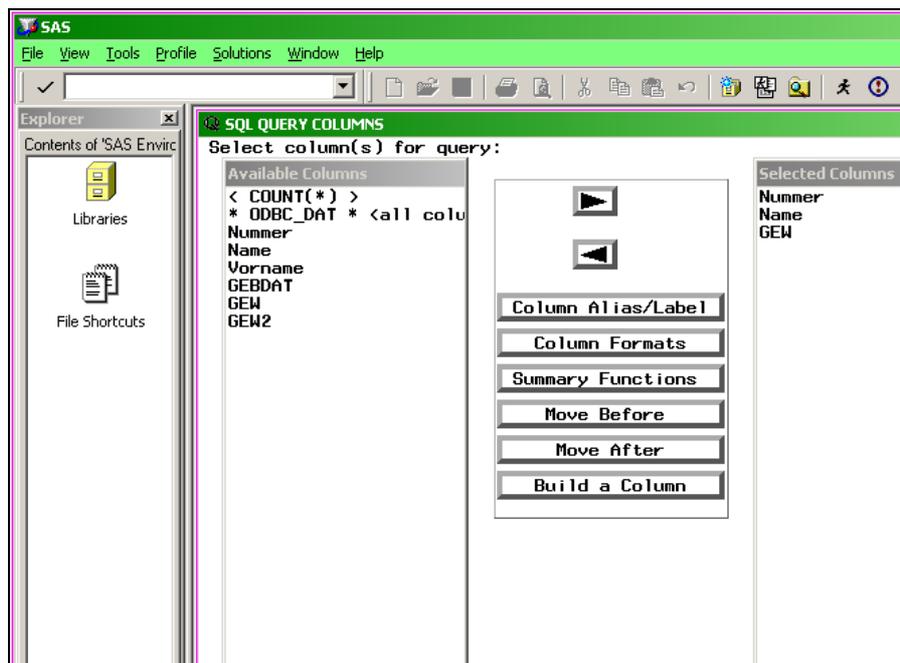


Abbildung 2.8: SQL Query Columns

Bemerkung 2.3:

Wurde vor dem Datentransfer über das SQL Query Window keine ODBC-Bibliothek definiert, gelangt man, falls man sich im SQL Query Window befindet über die Wahl von **Tools** → **Switch Access Mode** → **ODBC** und die anschließende Auswahl der entsprechenden ODBC-SQL-Datenquelle (`sql_dq`) ebenfalls zu den dort verfügbaren Tabellen.

3 Datentransfer zwischen SAS und Oracle

Da der Datentransfer zwischen SAS und Oracle im Prinzip analog zu dem in Kapitel 2 zwischen SAS und MS SQL Server beschriebenen verläuft, soll

hier nur kurz auf einige Oracle-spezifische Besonderheiten eingegangen werden. Dazu wird folgendes angenommen:

- Der (Windows 98) PC, auf dem SAS installiert ist, sei in einem lokalen Netz mit einem Server verbunden, auf dem Oracle8i läuft. Als Voraussetzung für einen Datentransfer muss auf diesem PC Oracle8i Client für Microsoft Windows installiert sein.
- Der Benutzer des SAS-PC hat mit dem Benutzernamen wodny und dem Passwort paul33 Lese- und Schreibzugriff auf eine Oracle-Datenbank mit dem Namen test (Server-Alias z.B. für die Definition einer ODBC-Oracle-Datenquelle), die eine Oracle-Liste mit dem Namen odbc_dat enthält (vgl. Abb. 2.1).

3.1 Datentransfer mit PROC SQL

Als Beispiel sollen wieder folgende Aufgaben gelöst werden:

- Erzeugen der temporären SAS-Datei data aus den Daten der Tabelle odbc_dat in der Oracle8i-Datenbank
- Ersetzen der in der Spalte GEW2 der Tabelle odbc_dat vorhandenen Werte durch die Werte 2*GEW

Das in Kapitel 2.1 aufgeführte SAS-Programm muss für den Datenaustausch mit Oracle wie folgt modifiziert werden:

```
/* 1. (ODBC-)Verbindung mit Oracle8i
   (im Intranet) über proc sql */
proc sql;
  connect to oracle ( user = wodny
                    pw   = paul33
                    path = 'test' ) ;

/* 2. Erstellen der (temporären) SAS-Datei data
   und Einfügen der Daten aus der Tabelle odbc_dat */
create table data as select * from connection
to oracle (select * from odbc_dat);

/* 3. Aktualisieren der in der Tabelle odbc_dat
   vorhandenen Spalte gew2 der Oracle-Datenbank */
execute (update odbc_dat
set gew2=2*gew) by oracle;

/* 4. Schließen der(ODBC-)Verbindung und Beenden von procsql */
disconnect from oracle;
quit;
```

3.2 ODBC-Datenquellendefinition für Oracle

Die ODBC-Datenquellendefinition für Oracle verläuft analog zu der im Kapitel 2.2 für SQL Server beschriebenen, allerdings wird jetzt der Oracle-Treiber gewählt. Ist die Software Oracle Client für Microsoft Windows nicht installiert, erscheint eine entsprechende Fehlermeldung und die Datenquellendefinition ist nicht möglich. Bleibt diese Fehlermeldung aus, erscheint das (hier englische) Fenster Microsoft ODBC for Oracle Setup (s. Abb. 3.1).

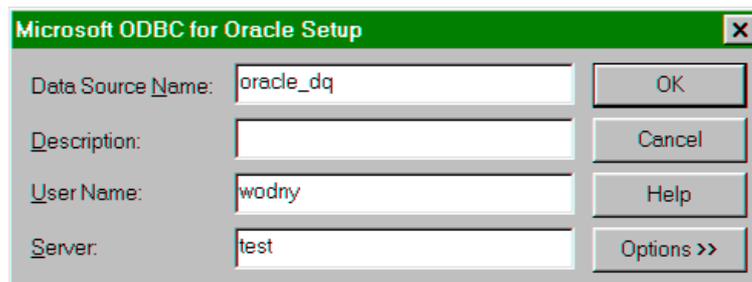


Abbildung 3.1: Microsoft ODBC for Oracle Setup

Nach der Vergabe des Datenquellennamens **oracle_dq** und der Eintragung von Benutzer- und Servername (Synonym für Datenbankzugriff) kann dieses Fenster mit **OK** verlassen werden und die Datenquellendefinition ist abgeschlossen.

3.3 Interaktive Bibliotheksdefinition

Die Interaktive Bibliotheksdefinition erfolgt ebenfalls analog zu der in Kapitel 2.3 beschriebenen. Die entsprechenden Abweichungen sind in Abb. 3.2 dargestellt.

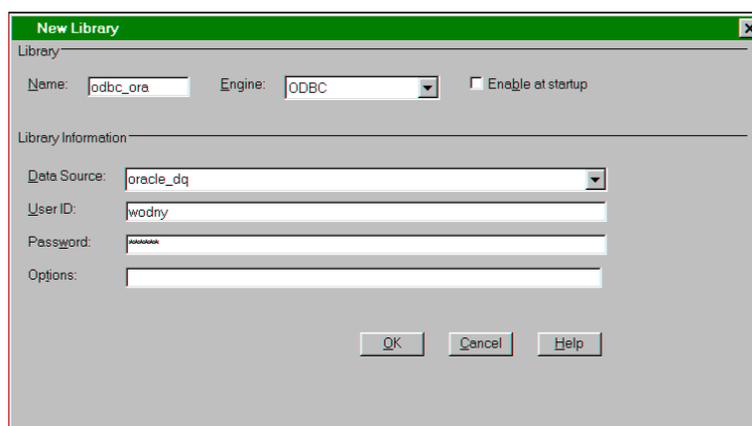


Abbildung 3.2: New Library

Das Analogon zum SAS/ACCESS LIBNAME Statement entsprechend Bemerkung 2.2 hat hier folgendes Aussehen:

```
libname ora odbc dsn = oracle_dq uid = wodny pwd = paul33;
```

3.4 Datentransfer mittels SAS SQL Query Window

Die Vorgehensweise beim Datentransfer ist hier vollkommen analog zu der in Kapitel 2.4 beschriebenen. Wurde vor dem Einstieg in den Transfer mit SAS SQL Query Window noch keine ODBC-Oracle-Bibliothek definiert, ist Bemerkung 2.3 entsprechend an dieser Stelle natürlich die ODBC-Oracle-Datenquelle (oracle_dq) zu wählen.

Literatur

- [1] Microsoft Corp. (1998). SQL Server-Onlinedokumentation, Version 7.
- [2] Oracle Corp. (2000). Oracle8i, On-Line Generic Documentation, Release 3 (8.1.7.).
- [3] Rudolph, P. E.; Geissler, H.; Ortseifen, C.(1998). Datenaustausch zwischen SAS und anderen Anwendungen mit DDE und ODBC in: Konferenzbeiträge der 2. Konferenz für SAS-Benutzer in Forschung und Entwicklung (KSFE), Friedrich-Schiller-Universität Jena, 200-220.
- [4] Rudolph, P. E.; Ortseifen, C.; Tuchscherer, A. (2000). DDE und ODBC zwischen SAS 6.12 und MS Office, in: Bödeker, R.-H.; Hollenhorst, M. (Hrsg.): Proceedings der 4. Konferenz der SAS-Anwender in Forschung und Entwicklung, 9.-10. März, Justus-Liebig-Univ. Gießen, 193-206.
- [5] Rudolph, P. E.; Ortseifen, C.; Havemann, P. (2001). Datenaustausch zwischen SAS 8 und anderen Windows-Applikationen mittels ODBC, in: Schumacher, E.; Streichfuss, K. (Hrsg.): Proceedings der 5. Konferenz der SAS-Anwender in Forschung und Entwicklung, 8.-9. März, Universität Hohenheim, 289-298.
- [6] SAS Institute Inc. (1999). SAS[®] Procedures Guide, Version 8, Cary, NC: SAS Institute Inc.
- [7] SAS Institute Inc. (1999). SAS[®] ODBC Driver User's Guide and Programmer's Reference, Version 8, Cary, NC: SAS Institute Inc.

- [8] SAS Institute Inc. (1999). SAS/ACCESS[®] Software for Relational Databases: Reference, Version 8 (ODBC Chapter), Cary, NC: SAS Institute Inc.
- [9] SAS Institute Inc. (1999). SAS SQL Query Window User's Guide, Version 8 Cary, NC: SAS Institute Inc.

