Introduction to Compustat North America (WRDS)

Authors:

Kenneth In Kyun Ernst Jørgensen

Last updated: **August 2008**

AARHUS
UNIVERSITY

Table of contents

1. Contents of the database

Compustat North America contains information on more than 24,000 active and inactive public companies. Compustat North America contains more detailed and specified information than Compustat Global and EMDB.

The information can be categorized within the following groups:

Company Financials

- Balance sheets and income statements
    - o Compustat Industrial Annually
    - o Compustat Industrial Quarterly
    - o Compustat Segment
    - o Compustat Index Fundamentals Annual
    - o Compustat Canadian Annual
    - o Compustat Bank Annual

Financial markets and prices

- Compustat Price, Dividends and Earnings

The information in the database is primarily accounting related, which e.g. can be used for valuation purposes. The advantage of the database is that it offers restated financials, which gives the opportunity for an easy comparison of the companies in the database.

2.1 Compustat Industrial Annually and Quarterly

Compustat Industrial is offered in an anually and quarterly version, which refers to the date range available. Industrial offers the most comprehensive search on specific company financial items. Full company financial statements are available under *Tools*, in which it is possible to extract complete financial statements or part of the balance sheet/income statement. The real strenght of Industrial is the decomposition of company financial items.

The decomposed items are not always available for the desired company, which depends on the reporting. An example of the grade of detail can be made with a comparison of the same search in Compustat and Datastream. In this scenario we want information for Novo Nordisk on Plant, Property and Equiptment (PP&E) for the balance sheet.

AARHUS
UNIVERSITY

Table 1: Datastream extraction of balance sheet item – Property, Plant and Equiptment – Net ($1000)

U:NVO	NOVO NORDISK A/S	
TYPE	DESCRIPTION	2006
		DETAILED
WC05350	DATE OF FISCAL YEAR END	31-12-2006
WC02501	PROPERTY, PLANT AND EQUIPMENT - NET	3599915

Table 2: Compustat North America Industrial Annually – Property, Plant and Equiptment ($MM)

Global Company Key	Data Date	Data Year – Fiscal	Industry Format	ISO Currency Code	Active/Inactive Status Marker	Property, Plant and Equipment – Total (Gross)	Property, Plant and Equipment – Total (Net)	Property, Plant, and Equipment – Ending Balance (Schedule V)
008020	20061231	2006	INDL	USD	A	5663.8750	3603.1090	5663.8750

A close look at the two searces indicates that Compustat offers both the Gross and Net value of PP&E for Novo Nordisk, while Datastream only provides the net value. In this case we have only extracted one element of the Asset side of the balance sheet. Several other items are available in Industrial with an equal or greater degree of detail.

2.2 Compustat Bank Annually

In accounting there is a difference between balance sheets and income statements for industrial companies and Banks. Compustat Industrial Annually and Quarterly represent both industrial companies and banks, where the banks financials are represented as Industrial statements. In addition to the information available on banks in Industrial, one can supplement with Compustat Bank annually, which features bank related items on the balance sheet and income statement.

2.3 Compustat Security Monthly

Security Monthly provides monthly historical data on the listed securities available in the database. The database contains opening and closing prices, adjustment factors, dividends, shares outstanding etc.

The advantage over e.g. Datastream is length of the data series available. For Security Monthly the earliest dated security prices are available from January 1962.

2.4 Compustat Pension

In the United States pension savings programmes are usually organized by the companies. U.S. companies are thereby obliged to report development in their pension capital. Compustat Pension offers unique information on this area. This kind of information is e.g. unavailable in Datastream. A relevant study could be a comparison analysis on pension performance among U.S. companies or a comparsion between U.S. companies and European Pension funds[1]. An example of statistics available are rate of returns, Pension benefits, Pension costs etc.

2.5 Compustat Industry speficic

Often we are not interested in specific companies, but more interested in specific industries. The most popular U.S. industries are categorized as Airlines, Homebuilding, Healtcare Facilities, Managed Healthcare (HMO), Oil and Gas, Retail and Utility. The use of these areas could be relevant for strategic analysis in a corporate valuation.

---

[1] Data for the European pension funds would have to be obtained from other sources.

AARHUS UNIVERSITY

3. How to get access?

Access to Compustat North America is part of the *Wharton Research Databases Services* and is available through the computers at ASB. WRDS is a database portal, which is accessible through http://lib.asb.dk/linkdb/. Besides Compustat North America the portal offers access to Compustat Global and EMDB, CRSP and NYSE TAQ.

4. How does it work?

The following example will illustrate a search on company financials for Novo Nordisk, which is listed on New York Stock Exchange, and thereby available in the database. The primary searching is quite alike in WRDS in general, and the search will thereby act as an example for further searches.

4.1 Overview of company codes and tickers

Searching in WRDS is available based on different ticker symbols and codes. The following sections will outline the difference and advantages of the different codes and tickers.

4.1.1 GVKEY

Compustat has defined a proprietary identifier, the GVKEY, for each company in the database. The GVKEY can be used to track a company over time, while the company name, CUSIP, or ticker may change over time.

4.1.2 SMBL

SMBL can usually, but not always, be identified with the ticker symbol of the company's publicly traded common stock. However, the same ticker can be assigned to different companies on different exchanges and company tickers can change over time. For example, the ticker "C" has been used by both Chrysler and Citigroup. Consequently, the ticker symbol is often an unreliable way to identify companies over time and in different databases.

4.1.3 CUSIP/CNUM/CIC

CUSIP is a nine-character alphanumeric identifier assigned to individual financial assets by an independent agency. As a rule, the first six characters of the CUSIP can be identified with a company; the next two characters (the seventh and eighth) identify a particular asset (e.g., a class of stock or a bond issue) issued by the company; and the ninth digit is a "check digit" to improve the accuracy of electronic transmission of CUSIPs.
Within Compustat, the first six characters of the CUSIP are referred to as the CNUM (CUSIP issuer code), and the last three characters of the CUSIP are called the CIC (CUSIP issue code). Throughout much of the history of Compustat data files, the CNUM was used as a company identifier, but it has recently been suplanted in this function by the GVKEY. In any event, the Compustat files do not track a company's CUSIP history. Rather, only the most recent CUSIP or CNUM is included.

4.1.4 DNUM/ NAICS

Until recently, Compustat used its DNUM variable to indicate a company's main line of business. This code was generally based on the SIC (Standard Industrial Classification) code created by the U.S. Census Bureau. The Census Bureau is replacing SIC codes with NAICS (North American Industry Classification System) codes. Further, Standard & Poor's and Morgan Stanley have developed GICS (Global Industry Classification System) codes.
A company's industrial classification can change from time to time, but the DNUM in Compustat data files generally reflects only the current classification, not the history of a company's industry affiliations.

4.2 Example of searching in Compustat North America

The startup in WRDS gives you the following options

By clicking on Compustat North America the following screen appears

4.3 Quick-Search: Fundamentals Annual

In our search we want information on the previous 5 years of financials for Novo Nordisk. By clicking on Fundamentals Annual, the following screen appears.

> Specify Fiscal Year and time period

> Specify the search: In this example we use CUSIP, but generally GVKEY is preferred, since it allows for unique tracking within

> The search yields the following results

The first part of our search specification is now completed. In the next step we need to specify, which parts of the financials we want to extract. In this example we will take a look at the information on specific *stock options* for Novo Nordisk.

The following search has to be specified

Data Items A-L
```
ACCHG -- Accounting Changes - Cumulative Effect
ACCLI -- Accrued Liabilities - Increase/(Decrease)
ACCO -- Acceptances Outstanding
ACDO -- Current Assets of Discontinued Operations
ACO -- Current Assets - Other - Total
ACODO -- Other Current Assets Excl Discontinued Operations
ACOMINC -- Accumulated Other Comprehensive Income (Loss)
ACOX -- Current Assets - Other - Sundry
ACOXAR -- Current Assets - Other - Total As Reported
ACQDISN -- Acquisitions and Disposals - Net Cash Flow
```

Data Items M-Z
```
OPTEXD -- Options - Exercised (-)
OPTFVGR -- Options - Fair Value of Options Granted
OPTGR -- Options - Granted
OPTLIFE -- Life of Options - Assumption (# yrs)
OPTOSBY -- Options Outstanding - Beg of Year
OPTOSEY -- Options Outstanding - End of Year
OPTPRCBY -- Options Outstanding Beg of Year - Price
OPTPRCCA -- Options Cancelled - Price
OPTPRCEX -- Options Exercised - Price
OPTPRCEY -- Options Outstanding End of Year - Price
```

Supplemental Data Items
```
ADJEX_C -- Cumulative Adjustment Factor by Ex-Date - Calendar
ADJEX_F -- Cumulative Adjustment Factor by Ex-Date - Fiscal
CSHTR_C -- Common Shares Traded - Annual - Calendar
CSHTR_F -- Common Shares Traded - Annual - Fiscal
DVPSP_C -- Dividends per Share - Pay Date - Calendar
DVPSP_F -- Dividends per Share - Pay Date - Fiscal
DVPSX_C -- Dividends per Share - Ex-Date - Calendar
DVPSX_F -- Dividends per Share - Ex-Date - Fiscal
MKVALT_C -- Market Value - Total - Calendar
MKVALT_F -- Market Value - Total - Fiscal
```

Under data items it is possible to specify, which financials to extract from the search. In our example we want information 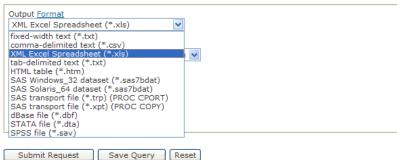on *Options*, which is selected from Data items M-Z. By holding down CTRL+ clicking on the specific items you can select the necessary information.

Step Four: Output

Output Format
```
XML Excel Spreadsheet (*.xls)
  fixed-width text (*.txt)
  comma-delimited text (*.csv)
  XML Excel Spreadsheet (*.xls)
  tab-delimited text (*.txt)
  HTML table (*.htm)
  SAS Windows_32 dataset (*.sas7bdat)
  SAS Solaris_64 dataset (*.sas7bdat)
  SAS transport file (*.trp) (PROC CPORT)
  SAS transport file (*.xpt) (PROC COPY)
  dBase file (*.dbf)
  STATA file (*.dta)
  SPSS file (*.sav)
```

Submit Request   Save Query   Reset

Under output format it is possible to specify, which data format to export to. In our example we will export to an Excel Spreadsheet. WRDS offers numerous output formats into the most popular statistical packages and text files.

By clicking submit request the following window appears

Data Request Summary

Data Request ID	200717581
Libraries/Data Sets	comp/funda /
Frequency/Date Range	ann / 2003 - 2007
Search Variable	CUSIP
Input Codes 2 item(s)	670100205
Conditional Statements	n/a
Output format/Compression	xls /
Variables Selected	CONSOL INDFMT DATAFMT POPSRC CURCD COSTAT OPTCA OPTDR OPTEX OPTEXD OPTFVGR OPTGR OPTLIFE OPTOSBY OPTOSEY OPTPRCBY OPTPRCCA OPTPRCEX OPTPRCEY OPTPRCGR OPTPRCWA OPTRFR OPTVOL
Extra Variables and Parameters Selected	C INDL STD

Your output is complete. Click on the link below to open the output file.

200717581.xls (13 KB, 4 observations 27 variables)

Download instructions
Netscape users... Shift-click
Internet Explorer users... Right-click and select "Save Target As..."

Right-click on the link, and choose save as. The spreadsheet can now be opened with the available data.

.
The resulting output looks as follows

Global Company Key	Data Date	Data Year - Fiscal	CUSIP	GIC Groups	Level of Consolidation - Company Annual Descriptor	Industry Format	Data Format	Population Source	ISO Currency Code	Active/Inactive Status Marker
008020	20031231	2003	670100205	3520	C	INDL	STD	D	USD	A
008020	20041231	2004	670100205	3520	C	INDL	STD	D	USD	A
008020	20051231	2005	670100205	3520	C	INDL	STD	D	USD	A
008020	20061231	2006	670100205	3520	C	INDL	STD	D	USD	A

Company Name	Ticker Symbol	CIK Number	Preliminary Date	Acquisitions and Disposals - Net Cash Flow	Options Exercisable (000)	Options - Exercised (-)	Options - Fair Value of Options Granted	Options - Granted
NOVO NORDISK A/S	NVO	0000353278	.	.	.	.	.	.
NOVO NORDISK A/S	NVO	0000353278	.	.	.	0.3950	.	0.8090
NOVO NORDISK A/S	NVO	0000353278	.	.	.	0.3900	.	0.9340
NOVO NORDISK A/S	NVO	0000353278	.	.	.	0.9530	.	1.1150

Options Outstanding - Beg of Year	Options Outstanding - End of Year	Options Outstanding Beg of Year - Price	Options Cancelled - Price	Options Exercised - Price	Options Outstanding End of Year - Price	Options Granted - Price	Options Exercisable - Weighted Avg Price	Risk Free Rate - Assumption (%)	Volatility - Assumption (%)
.	.	.	.	.	.	.	.	.	.
4.0380	4.4460	39.3200	39.3200	165.4500	41.3200	48.6000	.	.	.
4.4450	4.9760	36.0400	36.0400	30.0100	37.7900	48.5800	.	3.2500	15.0000
4.9760	5.0480	42.1400	42.1400	37.8900	47.4500	61.9700	.	3.6000	17.0000

5. Further resources

WRDS offers various manuals online that explains the different data types and how to export the data. Usually the manuals are available at the top of each category.

The following manuals are available online for Compustat North America.

THIS GUIDE HAS BEEN PRODUCED BY

ANALYTICS GROUP

Analytics Group, a division comprised of student instructors under AU IT, primarily offers support to researchers and employees.

Our field of competence is varied and covers questionnaire surveys, analyses and processing of collected data etc. AG also offers teaching assistance in a number of analytical resources such as SAS, SPSS and Excel by hosting courses organised by our student assistants. These courses are often an integrated part of the students' learning process regarding their specific academic area which ensures the coherence between these courses and the students' actual educational requirements.

In this respect, AG represents the main support division in matters of analytical software.

ADVANCED MULTIMEDIA GROUP

Advanced Multimedia Group is a division under AU IT supported by student instructors. Our primary objective is to convey knowledge to relevant user groups through manuals, courses and workshops.

Our course activities are mainly focused on MS Office, Adobe CS and CMS. Furthermore we engage in e-learning activities and auditive and visual communication of lectures and classes. AMG handles video assignments based on the recording, editing and distribution of lectures and we carry out a varied range of ad hoc assignments requested by employees.

In addition, AMG offers solutions regarding web development and we support students' and employees' daily use of typo3.

PLEASE ADDRESS QUESTIONS OR COMMENTS REGARDING THE CONTENTS OF THIS GUIDE TO

ANALYTICS@ASB.DK

AARHUS UNIVERSITY