Introduction to databases and research at the Finance Department

Mikko Niemenmaa
mikko.niemenmaa@aalto.fi
Chydenia G2.10
Three things I want you to remember after this session

• Read your research material before looking for data
• Plan enough time for the data retrieval process
• Don’t ask me questions about methods
Introduction

These slides should get you up-to-date on the types of data that are available at the Aalto University School of Economics Department of Finance.

This package also contains a primer on things to consider when starting the data related portion of your research and the actual retrieving and manipulation of your data.

Please refer to the material located at:

http://findb.aalto.fi/

for more information

In case of questions, omissions, and corrections, please contact:

Mikko Niemenmaa
Tel: 4313 8467
Chydenia G2.10
mikko.niemenmaa@aalto.fi
Motivation for learning about the data sources and how to use them

- Irrespective of the company or position that you will have in your future finance career, you will be required to retrieve data or at least know what data you want and where it can be found.
- Even if you have assistants who actually get the data, it is still beneficial to know what the process is and how long it takes.
- You might mention on your CV that you know how to use a particular database system.
Map of the location of my office in Chydenia on the 2nd floor
Map to the data resources in Chydenia on the 1st floor
Contents

• **Data usage process**
  • Retrieving data
  • Six examples
  • Performing analysis
The focus of these slides is on the data, analysis, and results part of a thesis.
The time taken by the data retrieval and analysis process is typically underestimated by students.

<table>
<thead>
<tr>
<th>Time taken</th>
<th>Design the analysis and experiments</th>
<th>Determine data requirements</th>
<th>Retrieve data</th>
<th>Set up data for analysis</th>
<th>Perform analysis and experiments</th>
<th>Create exhibits</th>
<th>Develop conclusions, write thesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Often, too little</td>
<td>Often, too little</td>
<td>Takes much longer than anticipated</td>
<td>Takes much longer than anticipated</td>
<td>Might take less time than expected</td>
<td>Often, too little time is left for this</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

How to prioritize your idea generating efforts for the research proposal

- Use judgment / intuition
- Involve your advisor
- Do back-of-the-envelope calculations
- Take risks

**Be practical!**

- Polishing:
  - Time and effort: 80%
  - Benefit for problem solving: 20%

- Focusing on impact:
  - Time and effort: 20%
  - Benefit for problem solving: 80%
Some ideas that seem good but that have underlying difficulties

<table>
<thead>
<tr>
<th>General research area idea</th>
<th>Potential challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mutual funds in Finland</td>
<td>• There is no single source of data for mutual funds outside of the US</td>
</tr>
<tr>
<td></td>
<td>• There are not that many mutual funds registered in Finland</td>
</tr>
<tr>
<td></td>
<td>• There are even less that invest only in Finland</td>
</tr>
<tr>
<td>Asset pricing in Europe</td>
<td>• There are no precalculated asset pricing factors for European markets</td>
</tr>
<tr>
<td></td>
<td>• It may be difficult to account for all the mergers and company specific events in different legal areas</td>
</tr>
<tr>
<td>High frequency returns of Finnish stocks</td>
<td>• The vastness of the data requires a special set of skills and technologies that takes time cultivate</td>
</tr>
<tr>
<td></td>
<td>• High frequency analysis in general requires a lot of data processing before results can be calculated</td>
</tr>
<tr>
<td>Chinese IPOs</td>
<td>• The availability of accurate information is limited</td>
</tr>
<tr>
<td></td>
<td>• The length of the timeseries is limited</td>
</tr>
</tbody>
</table>
Contents

• Data usage process

• **Retrieving data**
  Six examples

• Performing analysis
There is data available from the school, but remember to check and utilize other sources as well.

<table>
<thead>
<tr>
<th>Internal (HSE)</th>
<th>External</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Public</strong></td>
<td></td>
</tr>
</tbody>
</table>
| • E.g. Excel documents that are internal versions of public data | • Annual reports  
• Company web pages  
• Institute web pages  
• Government web pages  
• Press, public reports and articles |
| **Private**   |          |
| • HEX intraday data  
• Small firm accounting data | • Datastream  
• Thomson  
• SDC  
• BvD Orbis  
• Case company data |
### Often requested data and their sources

#### Market data

- **Stock Markets**
  - Datastream, Reuters EIKON, HEX, IBES
  - OMX / HEX
  - Daily price-,return- and intraday transaction databases
  - CRSP
- **Derivatives Markets**
  - Datastream Futures and Options Section
  - Reuters EIKON
- **Mutual Funds / Hedge Funds**
  - Lipper TASS
  - Domestic mutual funds
  - WRDS 13 F
  - Datastream and Reuters EIKON
  - International mutual funds

#### Fundamental data

- **International**
  - Thomson ONE / Worldscope
  - Orbis
- **Domestic**
  - Talouselämä 500 / ETLA
    - Financial statements from 1987
  - 500 largest companies in Finland
  - Voitto+
- **U.S.**
  - WRDS, Compustat

#### Event data

- **Corporate Finance (IPO, M&A, Venture Capital)**
  - SDC
- **News**
  - LexisNexis, Kauppalehti, Taloussanomat, Talouselämä, FT.com
  - OMX NewsClient
Data coverage of the various databases

<table>
<thead>
<tr>
<th>Market data</th>
<th>Fundamental data</th>
<th>Event data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thomson</td>
<td>I/B/E/S</td>
<td></td>
</tr>
<tr>
<td>Worldscope</td>
<td>Thomson</td>
<td>Research</td>
</tr>
<tr>
<td>Datastream</td>
<td>Reuters</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Voitto+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ETLA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Orbis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WRDS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(CRSP, Compustat)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OMX News</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lexis Nexis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SDC</td>
<td></td>
</tr>
</tbody>
</table>
Retrieving data is performed with a front-end program or an Excel add-in

**General features**

- Thomson ONE
  - Web based front-end for browsing and retrieving data
  - Excel add-in to facilitate retrieval of large data sets
  - Access available on two machines in the research lab

- Datastream Advance
  - Windows application front-end for retrieving data and web forms for browsing companies and data types
  - Excel add-in to facilitate retrieval of large data sets
  - Access available on one machine in the research lab

- SDC Platinum
  - Windows application front-end for retrieving data
  - Access available on one machine in the research lab

- Reuters EIKON
  - News, market data, and fundamentals are available from Reuters
  - A remote desktop connection is used to retrieve data
  - Saving to Excel is possible
  - Access available on five machines in the research lab
Retrieving data is performed with a front-end program or an Excel add-in.

**General features**

- **Web interface only**
- **Data on companies in the U.S. stock markets**
- **Various other data as well**

**WRDS:CRSP, COMPUSTAT**
- Web interface only
- Data on companies in the U.S. stock markets
- Various other data as well

**OMX NewsClient**
- Web interface only
- Stock market announcements and press releases from the OMX exchanges

**Orbis**
- Web interface only
- Accounting data and other firm information on worldwide small and large companies
- Access on all the computers on campus (and through the proxy)

**Excel and Access format**
- Some data sets are in Excel or Access format
- They are available on request
- They can be processed via formulas, SQL, or VBA (or any other language)
Contents

• Data usage process
• Retrieving data

Six examples

• Performing analysis
Six examples of data retrieval

1. **Reuters EIKON**
   - Logging on
   - Creating company list
   - Retrieving data in Excel

2. **Thomson ONE**
   - Logging on
   - Retrieving a company list

3. **Thomson ONE (Excel)**
   - Using the TF function

4. **Datastream**
   - Logging on
   - Retrieving a time series

5. **WRDS**
   - Logging on
   - Data available

6. **Voitto+**
   - Logging on
   - Saving data to a file
1

Reuters EIKON

- Logging on
- Creating a company list
- Retrieving data in Excel
Logging on to the Reuters machines

1) Log on to the workstation with the correct account:

   User name: Reuters1
   Password: **********

2) Double click the Reuters EIKON shortcut on the desktop
Logging on to the Reuters EIKON service

1) Click “Sign In”
The Reuters EIKON interface
Screening stocks in Thomson Reuters Eikon and retrieving a time series of fundamental data in Excel

1. Start up Screener app in Thomson Reuters Eikon

2. Create list of companies (make sure inactive ones are included)

3. Export list of companies to Excel

4. Use Build Formula wizard on the Thomson Reuters Ribbon tab to generate formula that retrieves time series data directly in to Excel
Start the Screener app in Thomson Reuters Eikon
Remember to include inactive companies in the screening universe
Pick the filtering criteria, e.g. the Country of Headquarters.
Look at the resulting list and Export it to Excel for further processing.
In Excel, start the Formula Builder by clicking the button.
Fill in the Formula Builder fields to get the data that you need.

```
=TR($A$2, "TR_TotalAssetsReported", SDate="20000101", EDate="20160101", CH=Date RH=Ed", E2)
```
Correct / Edit, copy-paste formula if necessary and retrieve required data

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Company Name</th>
<th>Country of Headquarters</th>
<th>Date</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMEAS.HE</td>
<td>Amer Sports Oyj</td>
<td>Finland</td>
<td>12/31/2000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STERV.HE</td>
<td>Stora Enso Oyj</td>
<td>Finland</td>
<td>12/31/2001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HUHV.HE</td>
<td>Huhtamaki Oyj</td>
<td>Finland</td>
<td>12/31/2002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KESBV.HE</td>
<td>Kesko Oyj</td>
<td>Finland</td>
<td>12/31/2003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>METSB.HE</td>
<td>Metsa Board Oyj</td>
<td>Finland</td>
<td>12/31/2000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OUTV.HE</td>
<td>Outokumpu Oyj</td>
<td>Finland</td>
<td>12/31/2001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STCBV.HE</td>
<td>Stockmann Oyj Abp</td>
<td>Finland</td>
<td>12/31/2002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YLEPS.HE</td>
<td>Yleislektronikka Oyj</td>
<td>Finland</td>
<td>12/31/2003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALBAV.HE</td>
<td>Alandsbanken Abp</td>
<td>Finland</td>
<td>12/31/2000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTYVS.HE</td>
<td>Citycon Oyj</td>
<td>Finland</td>
<td>12/31/2001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FISIV.HE</td>
<td>Fiskars Oyj Abp</td>
<td>Finland</td>
<td>12/31/2002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOKIA.HE</td>
<td>Nokia Corp</td>
<td>Finland</td>
<td>12/31/2003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLVAS.HE</td>
<td>Olvi Oyj</td>
<td>Finland</td>
<td>12/31/2000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>APETI.HE</td>
<td>Apetit Oyj</td>
<td>Finland</td>
<td>12/31/2001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YYYIV.HE</td>
<td>Yit Oyj</td>
<td>Finland</td>
<td>12/31/2002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEO1V.HE</td>
<td>Neo Industrial Oyj</td>
<td>Finland</td>
<td>12/31/2003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRA1V.HE</td>
<td>Cramo Oyj</td>
<td>Finland</td>
<td>12/31/2000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Thomson One Demo

- Logging on
- Creating a list of companies
- Retrieving data types
Logging on to a Thomson machine

User name: Thomson1
Password: **********
Open Internet Explorer and go to thomsonone.com

1. Enter the letter “s” in the username box and pick the username that becomes available.

   Click the “Sign In” button.

   DO NOT empty the IE cache, or no one else will be able to use the service.
Retrieving data for one company from Thomson ONE

Search for a company and browse the results by clicking on the different menu items.
Searching for a list of companies with Thomson ONE

1. Look for an exchange code, for example, to get a list of all currently listed stocks on an exchange.
2. Click the search button to get results.
3. You can also refine your search in the advanced search tab.
Results of a search

1. The number of results for your criteria are displayed.

2. Import the result list into excel to continue retrieving data for all the companies in the list.

3. You can also refine your current search by clicking the "Edit in Advanced Search" link.
Thomson One Banker (Excel) Demo

- Speed up processing: Remembering to turn off calculation and remove formulas when done.
- The TF formula
Turn off calculation to speed up processing
Copy and paste values to get rid of formulas

- Select the area you want to copy
- Press <ctrl>+c
- Paste special as values, Edit-> Paste Special
Using the TF formula to retrieve one value

=TF("NOK1V-HE","TF.RF.NetSales(period=FY2005)")

For which company
What you want
For what time period

34191000000

Result
Using the TF formula to retrieve a time series

\[ \text{TF("NOK1V-HE","TF.RF.NetSales(sdate=FY1990,edate=FY2011,frq=FY")}) \]

<table>
<thead>
<tr>
<th>For which company</th>
<th>What you want</th>
<th>For what time period</th>
</tr>
</thead>
<tbody>
<tr>
<td>303760000000</td>
<td>303760000000</td>
<td></td>
</tr>
<tr>
<td>311910000000</td>
<td>311910000000</td>
<td></td>
</tr>
<tr>
<td>300160000000</td>
<td>300160000000</td>
<td></td>
</tr>
<tr>
<td>294550000000</td>
<td>294550000000</td>
<td></td>
</tr>
<tr>
<td>292670000000</td>
<td>292670000000</td>
<td></td>
</tr>
<tr>
<td>341910000000</td>
<td>341910000000</td>
<td></td>
</tr>
</tbody>
</table>

Result
Be careful with the timing and frequency of your results when using the TF function.

This problem can be mitigated with the “Align Type” option in the formula editor.
Buttons you want to click and try in the Thomson Reuters Excel add-in

The Identifier lookup button lets you find company codes to use to retrieve data.

The Data Item Lookup button lets you create the formulas to retrieve data.

**WARNING!** May look different on your screen, but you can probably figure it out.
Datastream Demo

- Logging on
- Retrieving a time series
- Logging off
The Datastream interface and connecting to the service

• Click Tools -> Connect Now
Searching for companies

- Click the search button
Search results and selecting a company

- Click on the name of a company
Search for data types to retrieve

- Click the search button
Data retrieval results

<p>| | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Category:** Equities

**Search by:** Name

**Series:** 25687

**Databases:** Total Returns

**Request:** Request

**Project:** Project

**Equity Search:** Equity Search

**Request concluded:**
• Logging on
• Data available
WRDS is available everywhere!

Create your own username and password to get access.

Surf to:
https://wrds-web.wharton.upenn.edu/wrds/

1) Click to register an account

2) Fill out the form correctly. Use your advisor as the faculty contact.

Welcome to WRDS!

What is WRDS?
Wharton Research Data Services (WRDS) is a web-based b at the University of Pennsylvania. Developed in 1993 to su evolved to become a common tool for research for over 25

WRDS is the de facto standard for business data, providin economic, and marketing data though a uniform, web-bas locus for quantitative data research and is recognized by t the world as the leading business intelligence tool.

WRDS provides access to COMPUSTAT, CRSP, IBES, NYSE-Ti, other important business research databases.
## What is in WRDS?

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMPUSTAT North America</td>
<td>COMPUSTAT (from Standard &amp; Poor’s) provides more than 300 annual and 100 quarterly Income Statement, Balance Sheet, Statement of Cash Flows, and supplemental data items on more than 24,000 publicly held companies.</td>
</tr>
<tr>
<td>CRSP</td>
<td>The Center for Research in Security Prices maintains the most comprehensive collection of security price, return, and volume data for the NYSE, AMEX and Nasdaq stock markets. Additional CRSP files provide stock indices, beta- and cap-based portfolio, treasury bond and risk-free rates, and mutual fund databases.</td>
</tr>
<tr>
<td>I/B/E/S</td>
<td>Analyst information</td>
</tr>
<tr>
<td>Execucomp</td>
<td>Executive compensation</td>
</tr>
<tr>
<td>Audit Analytics</td>
<td>Auditor information</td>
</tr>
<tr>
<td>Bank Regulatory Blockholders</td>
<td>Fama French, Momentum, and Liquidity</td>
</tr>
<tr>
<td>CBOE Indexes</td>
<td>Federal Reserve Bank Reports</td>
</tr>
<tr>
<td>DMEF</td>
<td>PHLX</td>
</tr>
<tr>
<td>Dow Jones</td>
<td>Penn World Tables</td>
</tr>
<tr>
<td>FDIC</td>
<td>SEC Disclosure of Order Execution</td>
</tr>
<tr>
<td></td>
<td>TRACE</td>
</tr>
</tbody>
</table>
• Get a copy from me
• Exporting data
Voitto+ is available from me directly

Click on the Voitto+ icon. The following screen enables you to search for companies.
You can export data from Voitto+ into text files

1) Click “Select all” to select all of the companies in your search

<table>
<thead>
<tr>
<th>Company</th>
<th>Year</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub Nokia Oy</td>
<td>2007</td>
<td>140.0</td>
</tr>
<tr>
<td>Till-Nokia Oy</td>
<td>2007</td>
<td>329.0</td>
</tr>
</tbody>
</table>

2) Click the printer icon to “print” to a file

3) Select either financial or ratios to file (or csv)

4) Select the years to export, and the file location, then click to file.
Contents

- Data usage process
- Retrieving data
  Three examples
  - Performing analysis
The main software tools to perform data analysis

<table>
<thead>
<tr>
<th>Software Tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excel</td>
</tr>
<tr>
<td>Access</td>
</tr>
<tr>
<td>MATLAB</td>
</tr>
<tr>
<td>SAS</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>EViews</td>
</tr>
<tr>
<td>RATS</td>
</tr>
</tbody>
</table>

Alternatively, you can code / build your own in C, C++, VB, VBA, etc.
Free alternatives for statistical and mathematical software

<table>
<thead>
<tr>
<th>Software</th>
<th>Replaced by e.g.</th>
<th>Available from</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematica</td>
<td>Maxima</td>
<td><a href="http://maxima.sourceforge.net/">http://maxima.sourceforge.net/</a></td>
</tr>
<tr>
<td>Matlab</td>
<td>Octave</td>
<td><a href="http://www.gnu.org/software/octave/">http://www.gnu.org/software/octave/</a></td>
</tr>
<tr>
<td>SAS, RATS, SPSS</td>
<td>R</td>
<td><a href="http://www.r-project.org/">http://www.r-project.org/</a></td>
</tr>
<tr>
<td>EViews</td>
<td>gretl</td>
<td><a href="http://gretl.sourceforge.net/">http://gretl.sourceforge.net/</a></td>
</tr>
</tbody>
</table>
How do you typically set up your data for analysis?

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ret</td>
<td>YearID</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Company</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2.59</td>
<td>1.22</td>
<td>0.74</td>
<td>-0.18</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>0.26</td>
<td>0.10</td>
<td>0.51</td>
<td>1.45</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>0.92</td>
<td>3.05</td>
<td>-0.92</td>
<td>-0.79</td>
</tr>
</tbody>
</table>

Then it is easy to run an OLS regression:

\[
\text{Ret} = a + b_1 \times \text{Var1} + b_2 \times \text{Var2} + b_3 \times \text{Var3}
\]
Contents

• Data usage process
• Retrieving data
• Performing analysis

• Conclusion
The things I can help you with

• Using the databases
• Using different interfaces
• Using basic analysis tools
• Looking for data
Realize that there are few seats in the data room and they are in use by other students as well.
Next steps

1. Go to http://findb.aalto.fi
2. Fill out the form
3. Print it
4. Sign it
5. Return it to Niemenmaa's postbox at the finance department
END OF FILE