Jen Buckley, Vanessa Higgins and Jo Wathan
UK Data Service

… and the CESSDA Training Group
Overview

- Background: CESSDA and Data Discovery
- Training needs assessment
- Landscape mapping exercise
- Training activities
Consortium of European Social Science Data Archives

“to provide a full scale sustainable research infrastructure that enables the research community to conduct high-quality research which in turn leads to effective solutions to the major challenges facing society today”
National Data Services

Norway
NSD - Norwegian Centre for Research Data - CESSDA Member
Data discovery: a new area of expertise

help users locate and navigate data collections

a) relevant to their own research and teaching

a) in different locations and subject to different access conditions
Data discovery: 1\textsuperscript{st} Phase (2016-2017)

- Training needs assessment
- Landscape mapping exercise

Training
Training needs assessment

Aim: assess different training needs and training capacity Consult with national data services (time)

1. Survey of national data services
   - 22 individuals invite to online survey
   - response…
   - …19/22 (CESSDA members =14/15 and others=5/7)

2. Consult with representatives (from CESSDA Training Group)

3. Review of websites
Results: training important to users

- (k) Data citation
- (j) Undertaking cross-national analyses
- (l) Data visualisation
- (c) Evaluating for secondary data analysis
- (f) Survey weights
- (m) Teaching with data
- (i) Analysing data about geographical units
- (b) Finding and accessing: other sources
- (a) Finding and accessing: your service
- (e) Basic quantitative data manipulation
- (d) Exploring quantitative microdata
- (g) Complex samples
- (h) Analysing qualitative data

% agree

All
Members
Other

Results: training important to users
Results: some national data services provide training in data discovery topics

<table>
<thead>
<tr>
<th>Topic</th>
<th>Training events</th>
<th>Self-learning materials</th>
<th>Training events or self-learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Finding and accessing: your service</td>
<td>11</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>(b) Finding and accessing: other data sources</td>
<td>2</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>(k) Data citation</td>
<td>5</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>(d) Exploring quantitative microdata</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>(e) Basic quantitative data manipulation</td>
<td>3</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>(c) Evaluating for secondary data analysis</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>(m) Teaching with data</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>(f) Survey weights</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>(h) Analysing qualitative data</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>(j) Undertaking cross-national analyses</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>(l) Data visualisation</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>(g) Complex samples</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(i) Analysing data about geographical units</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

n=18
Finding Data at the ČSDA

ČSDA’s data are stored in the data catalogue based on the Nesstar system. Instruction how to use the catalogue can be found directly at the catalogue’s website. Bellow see instructions to help you obtain data from the ČSDA.

>>> ACCESS TO THE ČSDA CATALOGUE NESSTAR

>>> NESSTAR HELP AND SUPPORT

Modes for using the ČSDA data:

- **Free access without registration:** In the data catalogue, you can view the metadata (descriptive information) of each dataset deposited and you can display frequencies of...
About the UK Data Service
1,859 views  1 year ago
We provide access to the UK’s largest collection of UK and international social, economic and population data.
Poverty

Recently published datasets | Other datasets | Data series | Datasets elsewhere

FSD holdings contain a number of data series and individual datasets suitable for research on poverty. There are also many datasets that contain individual questions and variables which are potentially useful to researchers of poverty. The data archived at the FSD can be downloaded free of charge from Aila Data Service, according to the conditions set for each dataset. Please note that below is listed only a sample of all data available at the Archive and we encourage you to search more datasets on the subject on Aila.

Questions and variables of quantitative datasets that have been fielded in English or translated at the FSD can be explored by browsing their codebooks or through Aila variable search. If a quantitative dataset has not yet been translated, you can request a translation free of charge. Qualitative data are only available in their original language. Where applicable, the codebook and other related material are freely accessible on the study description page of each dataset.

Recently published datasets

<table>
<thead>
<tr>
<th>Data</th>
<th>Type</th>
<th>Availability</th>
<th>Data in English</th>
<th>Published</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSD3157 EVA Survey on Finnish Values and Attitudes 2017</td>
<td>quanti</td>
<td>B (1)</td>
<td>available on request</td>
<td>2017-04-05</td>
</tr>
<tr>
<td>FSD3133 Development Cooperation Survey 2016</td>
<td>quanti</td>
<td>B (1)</td>
<td>available on request</td>
<td>2017-04-03</td>
</tr>
</tbody>
</table>
Data analysis

Explore key variables, examine relationships between variables, data manipulation.

When planning and doing data analysis, there are several steps to consider. In this section, you will find information to help you:

- understand and explore key variables
- examine relationships between variables
- manipulate data to aid and improve analysis

There is a supporting Data Analysis Workbook with examples relating to the fear of crime and using the Crime Survey for England and Wales, 2013-2014.
Multinational Surveys

Several European countries participate in the following international continuous survey research programmes. Integrated and national data files are available either from the programme websites or through CESSDA archives.

Please follow the links below for further information about European and international surveys and how to access data.

Europe

European Commission Eurobarometer

Since 1973, the European Commission has been monitoring the evolution of public opinion in the Member States, thus helping the preparation of texts, decision-making and the evaluation of its work. Their surveys and studies address major topics concerning European citizenship: enlargement, social situation, health, culture, information technology, environment, the Euro, defence, etc.

Eurobarometer interactive search system
Context of training needs

- Some services provide training
- Training from other providers e.g. universities
- Remit of data services?
- Coverage of topics in secondary data analysis

Most of [this] training is given by universities and we do not know if there is a need for more training.

Almost [everything] is missing. There are of course various university courses that cover some of these topics, but nothing apart of that.

According to our experience, [...] training in data evaluation for secondary analysis is [an] important area still not met by courses and materials offered.
Results: Unmet training needs

- (m) Teaching with data: 84%
- (l) Data visualisation: 68%
- (j) Undertaking cross-national analyses: 68%
- (b) Finding and accessing data from other data sources: 68%
- (k) Data citation: 63%
- (h) Analysing qualitative data: 58%
- (g) Complex samples: 58%
- (f) Survey weights: 58%
- (c) Evaluating data for secondary data analysis: 58%
- (i) Analysing data about geographical units: 53%
- (a) Finding and accessing data in your service: 53%
- (e) Basic quantitative data manipulation: 47%
- (d) Exploring quantitative microdata: 37%

All
Other
CESSDA
Results: Unmet training needs

- Factors supporting competence in re-use of data – Finding data, evaluating data for secondary data analysis and data citation (“A, c and K are important to increase the competence level for enabling re-use of data”).

- We recognized a lack of knowledge related to data analysis and use of appropriate tools. Therefore data is used less that it could be.

- Providing more training of data discovery would enable researcher to find data they might use in their research projects without collecting data on their own.
Priority topics: topics important and where training is needed

<table>
<thead>
<tr>
<th>Topic</th>
<th>All (%)</th>
<th>Cessda (%)</th>
<th>Other (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data citation</td>
<td>63</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>Finding and accessing data: other data sources</td>
<td>58</td>
<td>50</td>
<td>80</td>
</tr>
<tr>
<td>Undertaking cross-national analyses</td>
<td>58</td>
<td>50</td>
<td>80</td>
</tr>
<tr>
<td>Data visualisation</td>
<td>58</td>
<td>50</td>
<td>80</td>
</tr>
<tr>
<td>Teaching with data</td>
<td>58</td>
<td>50</td>
<td>80</td>
</tr>
<tr>
<td>Evaluating for secondary data analysis</td>
<td>47</td>
<td>36</td>
<td>80</td>
</tr>
<tr>
<td>Survey weights</td>
<td>47</td>
<td>36</td>
<td>80</td>
</tr>
<tr>
<td>Finding and accessing data: your service</td>
<td>37</td>
<td>29</td>
<td>60</td>
</tr>
<tr>
<td>Basic quantitative data manipulation</td>
<td>32</td>
<td>29</td>
<td>40</td>
</tr>
<tr>
<td>Complex samples</td>
<td>32</td>
<td>14</td>
<td>80</td>
</tr>
<tr>
<td>Analysing qualitative data</td>
<td>32</td>
<td>36</td>
<td>20</td>
</tr>
<tr>
<td>Analysing data about geographical units</td>
<td>32</td>
<td>29</td>
<td>40</td>
</tr>
<tr>
<td>Exploring quantitative microdata</td>
<td>26</td>
<td>21</td>
<td>40</td>
</tr>
</tbody>
</table>
### Priority topics: insight from consultation

<table>
<thead>
<tr>
<th>Identified as important and needed by most by non-CESSDA services</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Finding and accessing: other sources</td>
</tr>
<tr>
<td>• Data Citation</td>
</tr>
<tr>
<td>• Undertaking cross-national analyses</td>
</tr>
<tr>
<td>• Data visualisation</td>
</tr>
<tr>
<td>• Teaching with data</td>
</tr>
<tr>
<td>• Evaluating for secondary data analysis</td>
</tr>
<tr>
<td>• Survey weights</td>
</tr>
<tr>
<td>• Complex samples</td>
</tr>
<tr>
<td>• Finding and evaluating data: using service</td>
</tr>
<tr>
<td>• Analysing qualitative data</td>
</tr>
<tr>
<td>• Analysing data about geographical units</td>
</tr>
<tr>
<td>• Exploring quantitative microdata</td>
</tr>
<tr>
<td>• Basic quantitative data manipulation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>- In progress</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Beyond remit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signpost</td>
</tr>
<tr>
<td>some elements</td>
</tr>
<tr>
<td>(e.g. missing data, complex datasets, metadata and documentation)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Share and repurpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signpost</td>
</tr>
<tr>
<td>some elements</td>
</tr>
<tr>
<td>(e.g. missing data, complex datasets, metadata and documentation)</td>
</tr>
</tbody>
</table>
Issues to consider when planning training

- Format – ‘local training’ through webinars and local workshops
- Adaptable materials – local training, local examples
- Sharing ideas and materials
- Statistics software (SPSS dominant)
Landscape mapping exercise

Aim: understand the researcher experience
What’s available? How to access?

1. review available data and access procedures for two topics:
   - ageing and political behavior

2. collate information about access conditions from national data services
Landscape mapping: some findings

1. Lots of data!
2. Separate catalogues/services (currently)
3. Searching is demanding (though tools to sort and refine)

Example search results (number of data collections)

<table>
<thead>
<tr>
<th>Data Collection</th>
<th>UK Data Service</th>
<th>GESIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Political behaviour</td>
<td>1098</td>
<td>3049</td>
</tr>
<tr>
<td>Political behavior</td>
<td>74</td>
<td>2979</td>
</tr>
<tr>
<td>voting</td>
<td>881</td>
<td>572</td>
</tr>
<tr>
<td>Online petition</td>
<td>26</td>
<td>633</td>
</tr>
<tr>
<td>Online AND petition</td>
<td>26</td>
<td>5843</td>
</tr>
<tr>
<td>Ageing</td>
<td>4887</td>
<td>12</td>
</tr>
<tr>
<td>Aging</td>
<td>4886</td>
<td>140</td>
</tr>
</tbody>
</table>
Landscape mapping: some findings

1. Lots of data!
2. Separate catalogues/services (currently)
3. Searching is demanding (though tools to sort and refine)
4. Language
5. NESSTAR – online analysis browsing (main catalogue or separate tool)
Data access

CESSDA general principle

available to anyone regardless of status, nation or type of use (except redistribution) unless there are known requirements which prevent it

But what about in practice?

• collated information from websites
• contacted services

Focus on the process for researchers outside of the country
Data access: common arrangements

Most often data collections can be downloaded by researchers (from any country) after simple registration

But also
- some open data
- apply for user name and password
- register use of data
- “Non-commercial use only” or “for research and teaching only”
- order data
- permission from data owner
- different arrangements for sensitive data
Training

Webinars
- How to find data in Europe (Wed 17 May)
- Data in Europe: Political Behaviour (Wed 7 Jun)
- Data in Europe: Ageing (Wed 14 Jun)

User guides
- Finding and accessing data from national data services
- Data discovery: Political behaviour
- Data discovery: Ageing

Finding and accessing data (video tutorial)
Available on CESSDA website – June/July
Training: a plan for the future

1. Workshop materials
   • finding and accessing data
   • missing data, complex datasets, metadata and documentation
   • CESSDA catalogue and question bank (due 2018)
   • available for use via CESSDA website

2. Share and signpost

3. Webinars
Conclusion

Issues
- role and remit
- diversity of needs
- reaching users

A way forward?
- signposting
- adaptable training materials
  - for local training
  - local examples
- webinars