

Jenny E. Symonds - jes81@cam.ac.uk

SECONDARY DATA ANALYSIS

Jenny Symonds
jes81@cam.ac.uk

What is it?



- The analysis of data from previous research
- Data is held in databanks by a range of providers
- The data is mostly free to use
- The analysis of old data can serve many purposes

Importantly, secondary data is:

- Available for **you** to use

Examples of data sets

Surveys and research projects

- British Cohort Studies
- Longitudinal Survey of Young People in England
- British Social Attitudes Survey
- Program for international student assessment (PISA)
- Trends in mathematics and science education study (TIMSS)

Statistics and other descriptive data

- National statistics data on class size, participation etc
- EduBase (data base of all English educational establishments)

What can I do with it?

Stand Alone Investigation

- Answer research questions -> publication/thesis

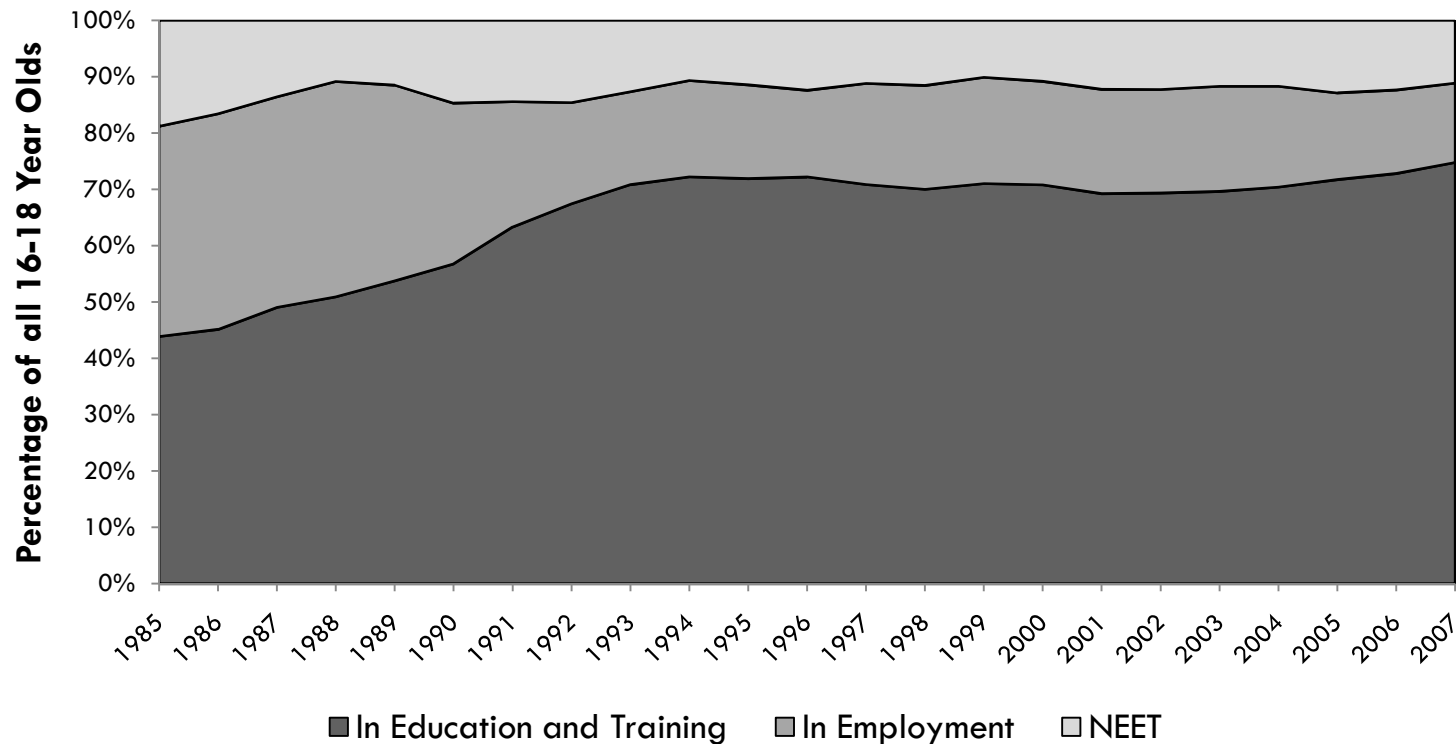
Within a Larger Investigation

E.g. MPhil or PhD thesis, research report

- Explore research interests -> literature review
- Confirm/confound theory -> literature review
- Inform data gathering -> pilot study
- Test findings -> discussion section
- Nested studies -> anywhere in document

Exploring research interests in a literature review

Percentage of 11-16 Year Olds in England in Education and Training, Employment, and NEET, 1985-2007



Symonds, Jenny E. - Compiled from data in the DCSF Statistical First Release on Participation in Education, Training and Employment (19th June 2008), additional information, excel files,

<http://www.dcsf.gov.uk/rsgateway/DB/SFR/s000792/index.shtml>.

Jenny E. Symonds - jes81@cam.ac.uk

Re-analysing old data for a new publication

Linnakylä & Malin (2008) Finnish Students' School Engagement Profiles in the Light of PISA 2003, *Scandinavian Journal of Educational Research*, 52:6, 583 -602.

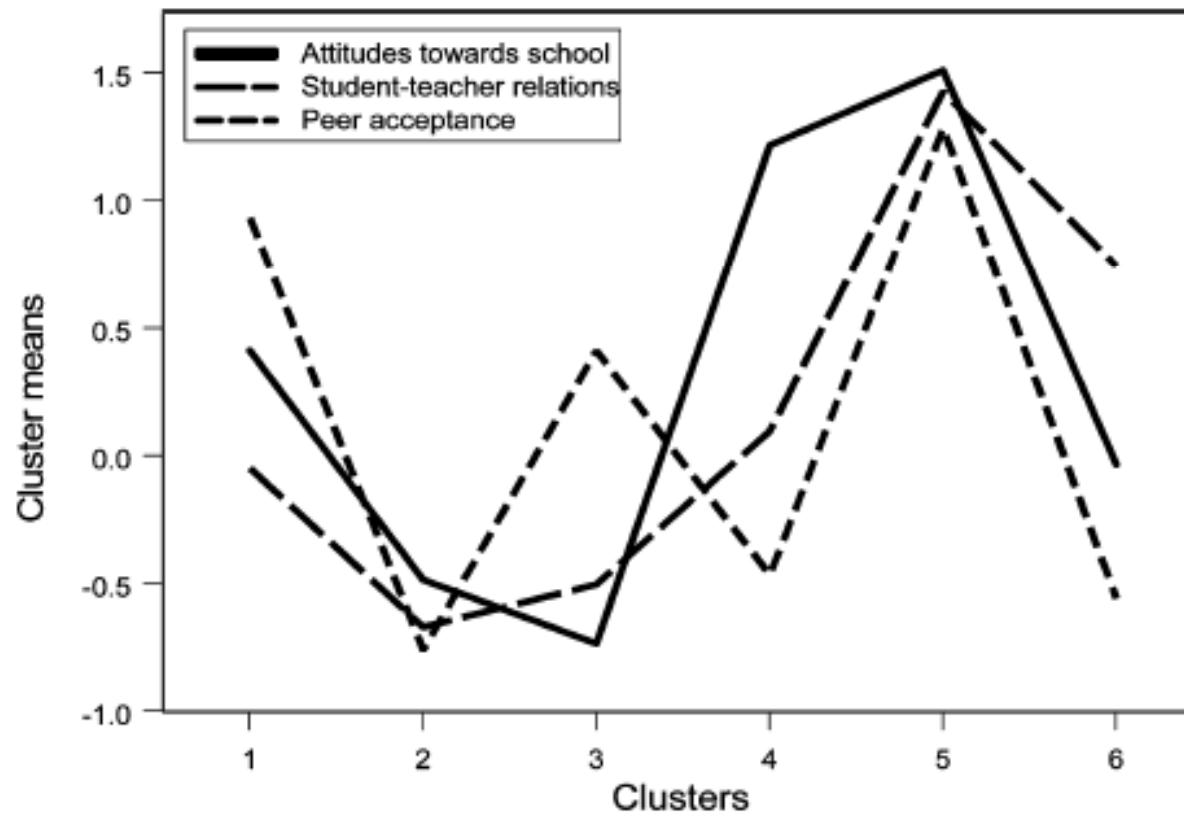


Figure 2. Results of cluster analysis based on the three engagement aspects

Jenny E. Symonds - jes81@cam.ac.uk

Creating compound variables for a macro-level analysis

UNICEF (2007) Child poverty in perspective: An overview of child well-being in rich countries, *Innocenti Report Card 7*, UNICEF Innocenti Research Centre, Florence

Dimensions of child well-being	Average ranking position (for all 6 dimensions)	Dimension 1	Dimension 2	Dimension 3	Dimension 4	Dimension 5	Dimension 6
		Material well-being	Health and safety	Educational well-being	Family and peer relationships	Behaviours and risks	Subjective well-being
Netherlands	4.2	10	2	6	3	3	1
Sweden	5.0	1	1	5	15	1	7
Denmark	7.2	4	4	8	9	6	12
Finland	7.5	3	3	4	17	7	11
Spain	8.0	12	6	15	8	5	2
Switzerland	8.3	5	9	1	4	12	6
Norway	8.7	2	8	1	10	13	8

Dimension made of:

PISA compound scores for English, maths and science
Country official statistics for % of 15-19 year olds in education

Dimension made of:

HBSC 'life satisfaction', 'liking school a lot' and 'health complaints' variables

A few words of warning...



- ❑ Downloading entire datasets can be challenging for the novice researcher
- ❑ Data often needs cleaning and sorting – this can take days if using large datasets
- ❑ When using large datasets (e.g. 14,000 cases), the researcher should be knowledgeable about the potential flaws in the data such as sampling error and the need to weight certain variables
- ❑ If in doubt – start small then build up later on

A few words of encouragement...



- Reanalysing data does not have to be tricky
- You can choose to construct a database with only a few variables for a few cases
- Often data is cleaned and sorted for you
- The analysis can be simple and non-time consuming (i.e. constructing a graph)

The simplicity of the task is up to you!

Where is the data kept?

□ Data Archives

- Economic and Social Data Service
- UK Data Archive (survey and 'quali-data')

□ Government Statistics

- UK Office for National Statistics
- US National Center for Education Statistics

□ Research Project Websites

- Program for International Student Assessment (PISA)
- Trends in International Mathematics and Science Study (TIMSS)

How can I find particular data?

□ Portals

- Special interest centres keep data bases of studies
- e.g. the Centre for Longitudinal Studies 'Keeping Track' database
<http://www.iser.essex.ac.uk/keeptrack/index.php>

□ Data Archives

- Lists of studies kept in the archive

□ Association Websites

- Often list studies of interest
- e.g. ISER, Longview, CAPCA, Centre for Longitudinal Studies

□ Internet Search Engines (general)

- e.g. look up project websites (PISA, PERLS, TIMSS)

□ Ask someone who is familiar with the field of study

How can I gain access to the data?

- Permission access
 - Email data manager/provider for access

- Declaration access
 - Sign up as a researcher to use the data
 - e.g. ESDS studies

- Free access
 - Should notify provider of any publications
 - e.g. PISA and TIMMS

Remember:

Secondary data analysis is a collaborative venture!

Exploring data before you download

- Some online services allow you to obtain basic descriptive data before downloading the entire data file
- Variables can sometimes be cross-tabulated
- Data can be used for preliminary analyses, if permissible

- Resources

ESDS: Nesstar Catalogue

<http://nesstar.esds.ac.uk/webview/index.jsp>

PISA Interactive Data Selection

<http://pisa2006.acer.edu.au/interactive.php>

Teaching with real data

- Using 'real data' in teaching is thought to engage students by giving them a 'real life' experience
- Teaching datasets are available from several providers

- Resources

ESDS learning and teaching with real data

<http://www.esds.ac.uk/findingdata/learning.asp>

Census at school

<http://www.censusatschool.org/>

Centre for longitudinal studies – teaching resources

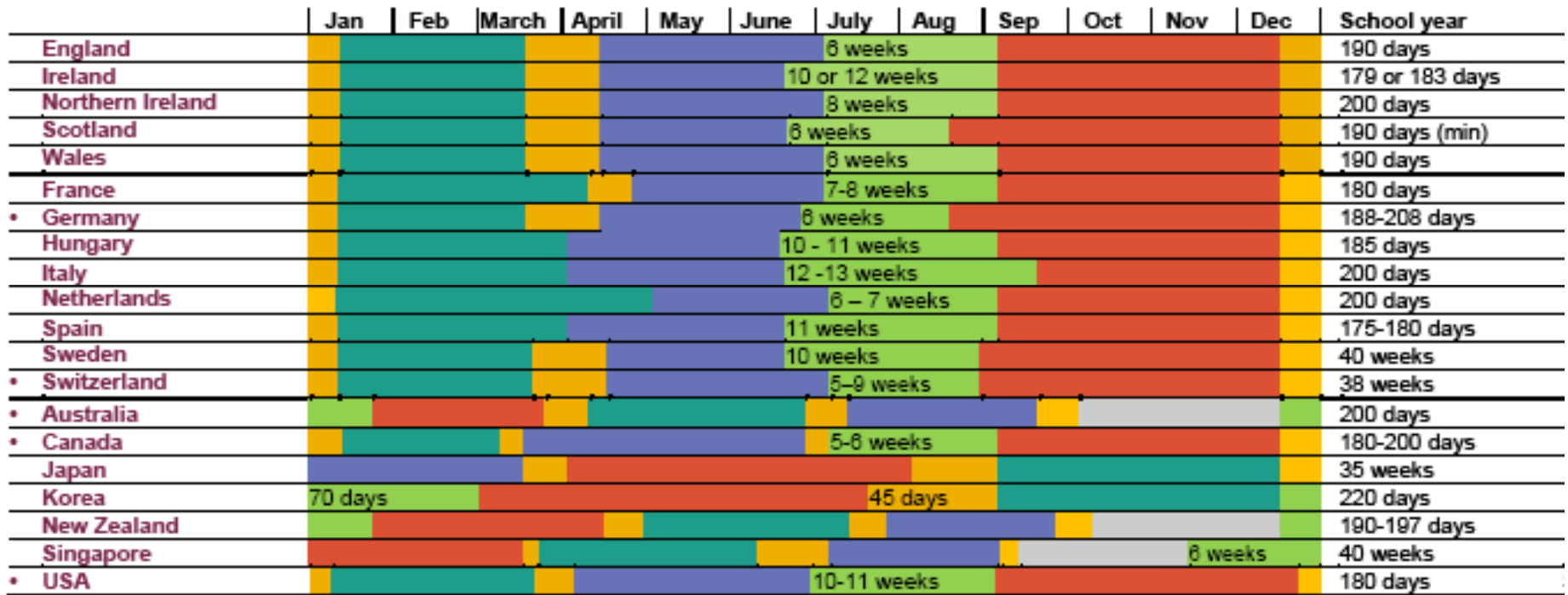
<http://www.cls.ioe.ac.uk/>

What if the data I need is not there?



- It is possible to be creative and make your own data sets by:
 - Making composite variables
 - Constructing variables using other data
 - Quantifying word based or categorical data

Example of transforming given data...



O'Donnell, Sharon & Sargent, Clare (2008) *Inca Comparative Tables*,
 Qualifications and Curriculum Authority

...into variables for a data set

School Structures

Number of long holidays	INCA 2008
Number of short holidays	INCA 2008
Number of terms	INCA 2008
Number of weeks of long holiday	INCA 2008
Number of weeks in school year*	INCA 2008
Number of weeks of short holidays	Computed from data

* Missing data for some countries computed from scores on other variables (e.g. number of weeks in the school year worked out by subtracting the number of weeks of holidays from the total number of weeks in a year) Jenny E. Symonds - jes81@cam.ac.uk

Example of categorical data for quantification

Classes of English Schools, abridged table, sourced from EduBase2 (Symonds, Jenny E.)

Key	Type of School	Maintained by State	Maintained by Private Sector	Selective Intake	Charges Fees
<i>Mainstream Schools</i>					
SS	Secondary School	Yes	No	No	No
MS	Middle School	Yes	No	No	No
HS	High School	Yes	No	No	No
Ac	Academy	Yes	Yes	No	No
CTC	City Technology College	Yes	Yes	No	No
SM	Secondary Modern	Yes	No	No	No
GS	Grammar School	Yes	No	Yes	No
ATS	All Through School	Yes	No	No	No
IS	Independent School	No	Yes	Yes & No	Yes & No

Some useful links

Lists of Data Archives

Emma Smith's website with links to the resources in her SDA book

- <http://www.secondarydataanalysis.com>

Research Resources for the Social Sciences (list of data archives)

- <http://socsciresearch.com/r6.html>

United Kingdom

UK Data Archive

- <http://www.data-archive.ac.uk/>

Economic and Social Data Archive

- <http://www.esds.ac.uk/>

Some useful links

International

Council of European Social Science Data Archives

- <http://www.CESSDA.org>

Inter-University Consortium for Political and Social Research

- <http://www.ICPSR.umich.edu>

Harvard Dataverse Network

- <http://dvn.iq.harvard.edu>

Other National

Australian Social Science Data Archive

- <http://ASSDA.anu.edu.au/>

Some useful links



Longitudinal Studies

Keeping Track - online search engine of longitudinal studies

- <http://www.iser.essex.ac.uk/keeptrack/index.php>

IOE Centre for Longitudinal Studies

- <http://www.cls.ioe.ac.uk/>

Longview - an independent think tank for longitudinal studies

- <http://www.longviewuk.com/>

Some useful links

Other Education Data

DfES EduBase2 - all registered educational establishments in the UK

- <http://www.edubase.gov.uk/home.xhtml>

DCSF and ONS – education statistics

- <http://www.dcsf.gov.uk/rsgateway/>

Learning and Skills Council – post-16 education statistics

- <http://www.lsc.gov.uk/providers/Data/statistics/>

National Pupil Database and Pupil Level Annual Schools Census

- <http://www.bristol.ac.uk/Depts/CMPO/PLUG/index.htm>

PISA 2006 data

- <http://pisa2006.acer.edu.au/downloads.php>

TIMSS

- <http://www.timss.org/>