Supporting Researchers

What do researchers need?
Services for researchers
What is administrative data?

- Information collected primarily for administrative/functional purposes

Source: Huntingdonshire District Council
Problems

• Multitude of datasets available with different procedures and approval bodies.
• Knowledge about admin data is limited and fragmented.
• Different attitudes towards release (DWP V. DFE).
• Different standards of quality and collection (local v. central).
• Non standardised variables for government departments
0930: Arrival, registration and coffee
1000: **Module 1** - An introduction to administrative data in the UK
   (Dr Chris Dibben, University of St Andrews)
1020: **Module 2** - Responsibilities of a safe researcher
   (Dr Mark Elliot, University of Manchester)
1110: Break
1125: **Module 3** - Creating safe settings and data security
   (Dr Mark Elliot)
1230: Lunch
1315: **Module 4** - Good practice in the analysis of administrative datasets
   (Dr Chris Dibben)
1415: Break
1430: **Module 5** - Understanding disclosure control and avoiding disclosive outputs
   (Dr Mark Elliot)
1600: **Module 6** – Making applications for administrative data
   (Dr Chris Dibben)
1630: Course close
Legal instruments

- Data Protection Act
- Other legislation
  - Breach of Confidence
  - The Human Rights Act 1998
  - Freedom of Information Act 2000
  - Computer Misuse Act 1990
  - The Statistics and Registration Services Act 2007
What is “personal data”?

☐ As we will see much privacy legislation uses the notion of ‘personal information’ or ‘personal data’

 But what is “personal data”?
 In pairs discuss this question and come up with 5 or 6 bullet points which may help answer this
Safe environments

- If you are wanting to have data to work on at your institution you will have to set up a safe setting for it.
- The nature of the safe setting will depend on the type, quantity, sensitivity etc of the data.
- Things you will need to consider include: software and hardware security, physical security, data and destruction transfer protocols.
- Explore available safe settings as alternative.
Good practice

- Ensure your security arrangements are sufficient.
- Only ask for data you actually need.
- Consider the need for ethical approval.
- Outputs – discuss with data holder how results will be published, used and to what level of detail.
- Contact the ADLS for further advice.
Some advantages of admin data

• Already collected for operational purposes
• Collection process not intrusive to target population
• Regularly updated
• Collected in a consistent way
• Near 100% coverage of population of interest
• Reliable at small area level
• Potential for datasets to be linked
Knife hospital admissions down 8%

The number of people being admitted to hospitals in England as a result of knife injuries fell by 8% in 2008, National Health Service figures show.

The findings also reveal that compared to the previous year there was a 14% fall in the number of teenagers admitted to hospital with stab wounds.

Meanwhile, figures from the Ministry of Justice suggest that more people are being jailed for possession of blades.

Police forces are using stop-and-search tactics to seize knives from youths.
Data linkage

Example 1

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HALL</td>
<td>STEPHEN</td>
<td>JOHN</td>
<td>MALE</td>
<td>22/07/38</td>
<td>14 HIGH STREET</td>
</tr>
<tr>
<td>HALL</td>
<td>STEPHEN</td>
<td>JOHN</td>
<td>MALE</td>
<td>22/07/38</td>
<td>14 HIGH STREET</td>
</tr>
</tbody>
</table>

Example 2

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HALL</td>
<td>STEPHEN</td>
<td>JOHN</td>
<td>MALE</td>
<td>22/07/38</td>
<td>14 HIGH STREET</td>
</tr>
<tr>
<td>HALL</td>
<td>STEVEN</td>
<td></td>
<td>MALE</td>
<td>22/08/38</td>
<td>14 HIGH STREET</td>
</tr>
</tbody>
</table>

MCS and health record linkage

- Admin data not obtained for all MCS respondents

<table>
<thead>
<tr>
<th></th>
<th>Birth registration data</th>
<th>Hospital records</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consent to match</td>
<td>90%</td>
<td>90%</td>
</tr>
<tr>
<td>Matching success</td>
<td>99%</td>
<td>83%</td>
</tr>
<tr>
<td>Total matched</td>
<td>89%</td>
<td>75%</td>
</tr>
</tbody>
</table>

- Some evidence of consent and match bias
Browse P-ADLS Resources

This page displays all resources currently available in the P-ADLS Bank. To view a list of the most recent resources added to the P-ADLS Bank click here.

Resources are organised by either computer code, research theme or universal classification codes and models. Computer code resources provide syntax and other files that can be downloaded and used.

- **Computer Codes (2 resources)**
- **Crime and Justice (4 resources)**
- **Education and Skills (25 resources)**
- **Health and Disability (13 resources)**
- **Housing (5 resources)**
The Disclosure Risk Problem: Type I: Identification

Identification file

<table>
<thead>
<tr>
<th>Name</th>
<th>Address</th>
<th>Sex</th>
<th>Age</th>
<th>..</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>Age</td>
<td>..</td>
<td>Income</td>
<td>..</td>
</tr>
<tr>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
</tbody>
</table>

Target file

ID variables

Key variables

Target variables
Action learning set

- In groups of four; pick one of your number who is considering setting up a safe data lab
  - That person discusses issues with setting up the lab
  - The others think about and provide feedback on the situation from other points of view:
    - The institutions
    - The data providing organisations
    - The respondents
    - A hacker!
Final observation