Big Data for Informed Decisions

ABS Big Data Strategy

Gemma Van Halderen, Population and Education Division, ABS
What is Big Data?

Rich data sets of such size, complexity and volatility that it is not feasible to fully leverage their information value with existing data capture, storage, processing, analysis and management practices.
The Internet of Everything

The emerging “network of networks” linking together people (the social web), information (the traditional Web), things (the sensor web), and places (the geospatial web)

These interconnected networks continually produce and consume data
Sources of Big Data

Digital descriptions of the physical environment (geography, geology, buildings, maps, environment, weather)

Sensors and other devices (GPS, asset tags, phones, flow meters, temperature sensors, medical instruments)

Logging and tracking of individual behaviour and the information they create and use (email, tweets, social network interactions, phone calls, web page visits and clicks)

Digitisation of commerce and supply chains (asset movements, orders, inventory, payments), financial assets and transactions
The ABS Vision for Big Data

Harness diverse sources of Big Data to

- Create a richer, more dynamic and better focused statistical picture of Australia to meet emerging needs
- Reduce the cost of statistical production and support
- Improve the relevance and timeliness of statistics
Enabling the ABS Vision

A skilled workforce able to interpret information needs and communicate the insights gleaned from rich data

Advanced methods, tools and infrastructure to represent, store, manipulate, integrate and analyse large, complex data sets

A diverse pool of government, private and open data sources available for statistical purposes

Safe and appropriate public access to microdata sets and statistical solutions derived from an array of data sources

Strong multidisciplinary partnerships across government, industry, academia and the statistical community.
# Specific Strategies

<table>
<thead>
<tr>
<th>A skilled workforce able to interpret information needs and communicate the insights gleaned from rich data</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Build and share competency in data science</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Advanced methods, tools and infrastructure to represent, store, manipulate, integrate and analyse complex data</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Enhance data integration methods, tools and infrastructure</td>
</tr>
<tr>
<td>• Modernise statistical practice for non-traditional data sources</td>
</tr>
<tr>
<td>• Introduce new approaches to data modelling and analysis</td>
</tr>
<tr>
<td>• Evaluate and deploy high performance computing platforms</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A diverse pool of government, private and open data sources available for statistical purposes</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Formalise the benefit assessment framework for Big Data</td>
</tr>
<tr>
<td>• Facilitate the sharing of private data for public good</td>
</tr>
<tr>
<td>• Trial the targeted use of external data provision services</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Safe and appropriate public access to microdata sets and statistical solutions derived from an array of data sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Lead the national adoption of privacy preserving data analytics</td>
</tr>
<tr>
<td>• Develop microdata access solutions with strong built-in confidentialisation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strong multidisciplinary partnerships across government, industry, academia and the statistical community</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Support and leverage external system development initiatives</td>
</tr>
<tr>
<td>• Establish a broad network of research collaborators</td>
</tr>
</tbody>
</table>
Big Data, Big Challenges

- Business benefit
- Privacy and public trust
- Methodological soundness
- Technological feasibility
- Data acquisition
Big Data R&D Flagship Project

Build a strong foundation for the mainstream use of Big Data in statistical production

- Methods
- Skills
- Tools and infrastructure

Through a coordinated set of targeted R&D initiatives

- Match Big Data opportunities to specific business problems
- Deliver “fit for purpose” solutions as working prototypes
- Enhance partnerships with academia, industry and other NSIs
- Contribute to a whole-of-government capability
ABS Research Areas

Satellite and ground sensor data for agricultural statistics
Mobile positioning data for measuring population mobility
Financial transactions data for macroeconomic statistics
Scanner and other point-of-sale data for prices statistics

Predictive modelling of survey non-response behaviour
Predictive modelling of unemployment for small areas
Data visualisation techniques for exploring large datasets
ABS Strategic Priority

ABS Executive Leadership Group have endorsed the strategy

Ministerial support for ABS to do something
  - but do it sooner rather than later)

Strategic priority to get on with demonstrating what can be delivered with the available data
  - not what can’t be delivered because of data limitations
Thank you