
Making New South Wales Religion, Education and Population Statistics Accessible

Crichton Smith,
Nick Parr,
Nikola Balnave,
Lucy Taksa
and
Brian Croke

Introduction

The website www.REPstats.info provides free on-line access to explore and graph complete time series statistics from 1828 to the present day on religion, education and population for the former British colony now state—of New South Wales (Australia). Drawing from the New South Wales (NSW) comprehensive collection of official statistics, the website brings nineteenth century data into twenty-first century formats, allowing researchers and other interested users to access the information as graphs and tables or drop the data into Microsoft Excel and other software. REPstats (Religion, Education and Population Statistics) has brought the tables of data from the pages of annual Colonial Statistical Registers, Population Censuses and other official statistics to facilitate research. Users can download the original tables for specific years as Excel files, complete with technical notes, or see a graph for all years with just a few clicks of the mouse, placing the statistics in historical and contemporary contexts.

Early statistics

European settlement in NSW commenced as a penal colony in 1788 with the arrival of the first fleet. In 1828, NSW replaced population musters with its first census, an enumeration of the population and its characteristics, including age, sex, year of arrival, occupation and religion. The second census in 1841 expanded to include marital status and civil status (free and bond, natives and immigrants), housing and education. A detailed (144-age) analytical view of the Census reported on the colony and, for Sydney and the rest of the colony, gave details of the numbers of males and females by age who ‘Cannot Read’, can ‘Read only’, or who could ‘Read and Write’.¹ Subsequent censuses expanded the analysis to cross-tabulate literacy by religion, birthplace, and those under tuition at home or school.

1 R. Mansfield, *Analytical View of the Census of New South Wales for the Year 1846* (Sydney, 1847).

In addition to censuses, the NSW Legislative Assembly commenced an annual collection and publication of reports from the Registrar General in 1857. The first annual report was 19 pages and reported births, marriages (by 11 denominations in Sydney and country), and deaths. The report was particularly focused on deaths in the colony, with 10 pages devoted to tables enumerating causes of death by age. Within a decade the statistical report had expanded to over 250 pages, including statistics on topics as diverse as university enrolments, Sunday school attendance, crime, industry, and climate. By the 1890s, the Statistical Register was 600 pages and provided a comprehensive annual statistical understanding of the colony.

The REPstats project

The volume and detail of these statistics provide researchers with a richness of data, but access to the reports has undoubtedly been a barrier to their usage. Few hard copies of early statistics remain, and access to them is generally restricted. Furthermore, transcribing data from individual reports is time-consuming with a tendency for researcher to duplicate the task.

Fortuitously, in 2010 we were able to obtain 90 boxes of statistical publications, including the almost complete collection of NSW Statistical Registers, colonial censuses, and national censuses (from 1911) from the Australian Bureau of Statistics. In July 2015, the Catholic Education Commission NSW (CECNSW) partnered with the Faculty of Business and Economics at Macquarie University, financially and intellectually, to create an opportunity for the data to be liberated from the original reports by creating an accessible electronic historical repository of religion, education and population. Within 17 months, REPstats was released with over 250 Excel files relating to education, religion and population to allow researchers to access data for individual years from 1821 to 1971. REPstats also released the first accessible longitudinal dataset (religion by sex for NSW), allowing users to see a graph and table of the entire composition of religions and denominations reported for NSW from 1828 to 2011.

The REPstats approach

REPStats was conceptualised to enable researchers to access and explore time series of data, in order better to identify trends and/or unusual events. In doing so, the project provides users with Excel files for individual years of the Statistical Registrar and Census statistical tables. While time-consuming to transcribe, Excel files for individual tables provide greater contextual information of the table structure, classifications of the data, and nomenclature than graphs allow. These files include the technical notes from the original reports and notes from the researchers transcribing the information.

REPstats also allows researchers quickly to access data for all years in a single table. With time series statistical data, complications arise from changes in definitions and classifications. The intent behind REPstats is to provide the data, rather than determine the

appropriateness of including or excluding data. To achieve this, the time series datasets present a complete list of categories, even though some of these categories may be similar or archaic, maintaining the historical integrity of the data. A feature of REPstats is that, rather than users being deterred by the multitude of statistical categories over time, the website allows users to access higher level aggregates of the classifications, as defined by the statistical classifications of the day, rather than according to contemporary classifications. For example, Catholics have been a religious category in every census, but there have been 31 different sub-classifications throughout the 27 censuses since 1828. Greek Orthodox were included as Catholic until the 1954 Census, when they were recognised as not being a Catholic denomination (the changing categorisation perhaps reflecting changing attitudes and recognition as a result of post-war migration).

Maintaining the nomenclature of the day results in multiple categories, but provides researchers with new insights into the social and statistical construct of NSW society over time. For instance, the contemporary statistical category of 'Islam', since first reported in 1851—as 'Mahomedans and Pagans'—has been reported by eight categories including 'Mahometan', 'Mohammedan' and, more recently, 'Muslim'. 'Hebrews' (to use the term employed by the largest number of censuses) has 11 categories including 'Jews', 'Jewish', 'Judaism', 'Israelite' and 'Synagogue'. The shifting nomenclature from 'Jews' to 'Jewish' to 'Judaism' offers insight into the evolution of Australia's statistical classifications from the person ('Jews') to the statistical/technical ('Judaism'). Similarly, the 273 Christian Sects, and 22 categories of Freethinkers—in the three censuses from 1891 to 1911—provide insight into the shift in statistical practices, and the religious and social constructs at the birth of Australia's federation.

Using the website

Users visiting the website www.REPstats.info have several options at the landing page.

The screenshot shows the REPstats website landing page. At the top, the title "REPstats" is displayed in a large, bold, sans-serif font. Below the title, a subtitle reads: "The Catholic Education Commission NSW and Macquarie University REPstats is an exciting on-going project making NSW statistical data available for the first time in an easy to access format." The main content area is divided into five dark grey rectangular boxes, each with a white icon and text. The first box, titled "About", features a circular icon with a cross and a church steeple, and contains the text "Read about the ideas and the team behind REPstats". The second box, titled "Available Data", shows a stack of books and says "Find out what type of data is available". The third box, titled "Get Some Data", displays a bar chart and says "Do a query and show tables or graphs with the results". The fourth box, titled "Download Files", shows a document titled "STATISTICAL REGISTER New South Wales 1882" and says "Download the original data as Excel worksheets". The fifth box, titled "Sign Up", features an envelope icon with an '@' symbol and says "Sign up to get news as REPstats is updated". At the bottom left, there is a small line of text: "A project by the Catholic Education Commission NSW and Macquarie University. Visit the Australian Catholic Education Statistics Portal."

Those interested in accessing the Excel files for individual years can click the <Download Files> button, while users interested in exploring longitudinal data click the <Get Some Data> button. Hopefully, some of you will choose <Sign up> so that we can keep you updated on REPstats developments over the coming months.

When you select <Get Some Data>, REPstats takes you through a straightforward process for selecting the data and output, visualised by the following steps (which you can click on and skip steps):

- <Choose a Topic>
- <Choose a Geography>
- <Items>
- <Limits>
- <Get Results>
- <New Query>

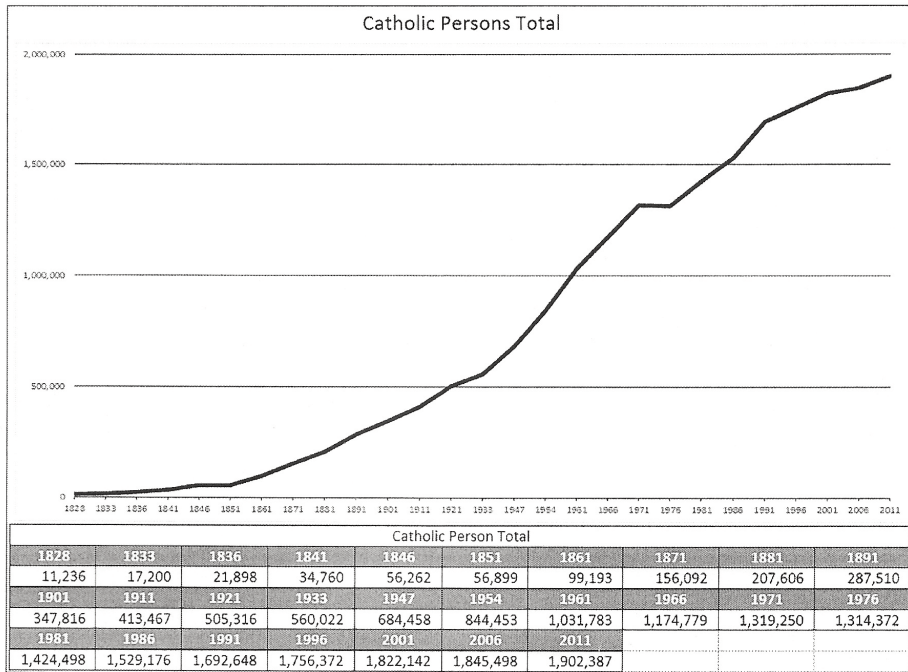
At this stage, REPstats has only one dataset (this will expand over the next few months). So, using this dataset as an example, the process is:

- <Choose a topic> Religion
- <Choose a Geography> NSW Total
- <Select the Data> Choose the item(s) you want to display (for example Catholic Totals—Persons)
- <Set some limits> You can set time for the beginning and end of the display. Leave it at the default of <All years>
- <Get the Data> You can choose various types of chart to display the data. Choose a <Line Chart>.

The screen capture below shows the outcome: both a graph and table are produced. At this point, you can print or download the chart, copy and paste the table into Excel, or go back and explore the data further.

Where next?

REPstats is a work in progress. Over the next few months, the researchers will be converting much of the information in the annual files into datasets to expand the number of time series datasets, and include datasets relating to population and education, as well as additional religion datasets. If you would like to suggest topics that interest you in the downloadable files to be converted into time series datasets (allowing the data to be graphed), or suggest other future directions for the project, please email Crichton.smith@hdr.mq.edu.au or phone (+61) 0410 544 157 to discuss. We look forward to your involvement in developing REPstats.



Acknowledgements

We would like to acknowledge the financial contribution from the Catholic Education Commission, NSW and Macquarie University through the Macquarie University's Enterprise Partnership Scheme.

We wish to acknowledge the contribution for our research assistants Sheruni De Alwis, Stephanie Hull, Victoria Morley and Virginia Mandelburger, and our website developer Scott McDonald. REPstats would not be possible without their efforts.