“From a Publishers’ Perspective”

Creating a large-scale collection of genuinely teacher-ready teaching datasets

Bronia Flett & Patrick Brindle
4th June 2014
Overview

- SAGE Research Methods
- Ideas and the concept of SAGE Research Methods
- Datasets
- Idea Development: Identifying the need and market research, how to add value to what exists out there?
- Planning and commissioning
- Challenges and opportunities

Questions/Comments greatly appreciated!!
http://srmo.sagepub.com/
175,000 pages of SAGE book, journal and reference content
Methodological focus so cross disciplinary
Advanced search and discovery tools, Methods Map.

http://srmo.sagepub.com/cases
Where do ideas come from?

- **Locating a need**
  - What problem can we solve for our community?
- **Understanding the market**
  - Teachers of research methods
  - Postgrad researchers
  - U/g students
The Idea: SAGE Research Methods Datasets

- Identifiable need for Quantitative datasets
- Difficulties in sourcing and cleaning
- Pedagogical AND interesting
Idea Development: Assumption or a real need?

- **Market research**
  - 60 academics completed a survey on datasets idea (all 60 in favour)
  - Backed up by a major survey on SAGE’s RM strategy involving 2,000+ respondents (librarians & academics)
What the market research told us

- Product solves a big problem for faculty
- Product idea something that librarians would support buying
- Need is for teaching datasets (to support RM) more than substantive research data
- Many teaching datasets already available but hard to use in practice
- Need to add value to the datasets to make them more ‘classroom ready’ than existing online resources
What do faculty need?

- Genuinely ‘teaching ready’ datasets
  - Absence in the market, despite a seeming profusion
  - Lots of datasets but so hard to use for beginners and hard to use for teachers too
  - Datasets that support lab work and individual learning
  - Major finding: need for Qualitative Datasets too
    - Ongoing research to shape the qualitative offering
Summary of important points

- Datasets need to be
  - Interesting (often sadly overlooked)
  - Discoverable
  - Meaningful to students
  - A good fit to core disciplines
  - Downloadable to LMS or VLE
  - Fully supported by teaching notes
  - Up to date
  - International
  - Based on real research data
What should SAGE Datasets look like?

- **Adding value through development**
  - Supporting narratives
  - Meta-data and Bibliographical information
  - Visualisations
  - Charismatic and interesting data
  - Code books and instructions on use
  - Global mix
  - Discipline mix
  - Datasets to be downloadable in a variety of formats to support lab work
    - SPSS, R, Excel, Stata and .csv formats
Planning and Commissioning

- **Quantitative Content provider: Who?**
  - Odum Institute, University of North Carolina

- **Sourcing datasets**
  - International (40% US focused/located)

- **What methods? What disciplines?**
  - 14 introductory methods across 6 disciplines (84!)
  - 14 advanced methods across 2-3 disciplines (21)
Challenges and Opportunities

● **Qualitative**
  • How do you develop datasets for QDA?
  • Cardiff University – data providers and partners in this challenge!

● **Integration with SRM**
  • How “tied” to SRM should Datasets be?

● **Sales**
**The 2007 School Readiness Survey**

**Lee Williamson, Mark Brown, Jo Watthee, Vanessa Higgins**
National Center for Education Statistics in the United States Department of Education

**Date of Collection**: September, 2007 - January, 2008 | **Data Production Date**: 2014

**DOI**: http://dx.doi.org/10.31938/16420649

**Tags**: Logit models | Logistic regression

**ABSTRACT**

The National Center for Education Statistics in the United States Department of Education operates a National Household Education Surveys Program that conducts a range of surveys regarding educational activities in the United States. Various studies cover the educational activities of people of all ages, including early childhood through adulthood. The most recent survey was conducted in 2012, but publicly available data sets can be found online from 1991 through 2007.

This example uses data from the School Readiness Survey. [DOI GOES HERE] was conducted in 2007 designed to gather information on a range of activities and behaviors for children ages 3 through 6 who were not yet enrolled in kindergarten. The survey also collected information on characteristics of each child’s parents or formal guardians. The full survey includes 940 variables. The size, scope, and complexity of the 2007 School Readiness Survey cannot be described in detail here. In particular, detailing the sampling design and the proper use of weighting in any analysis is beyond the scope of this document. Users are STRONGLY encouraged to read the full documentation for the 2007 School Readiness Survey before using it for research. [SAGE related content goes here]

**HOW-TO GUIDE**

**Dependent Variables**

Logit models are designed for dependent variables that have two mutually exclusive and exhaustive categories. Such dichotomous variables are often coded as either “1” or “0”, with a 1 indicating the presence of some attribute or behavior and a 0 indicating its absence. In this example, we use four dependent variables.

**Related Datasets**

Related datasets can be found here.
What next?

- Add content, features and functionality to datasets collection

- Add further enhancements to SRM
  - Video
  - Classic works?
  - Project planner
Questions and comments are very welcome!

Bronia.flett@sagepub.co.uk