



Dataset Review—"Data-Planet: A Successful Private-Sector Data Resource"

Repository: <https://data-planet.com>

Author/Compiler: Richard Landry,

Director of Data Innovation at SAGE Publishing and CEO of Conquest Systems, Inc.

Reviewed by Patrick Manning

University of Pittsburgh

pmanning@pitt.edu

The Data-Planet website, centering on its Data-Planet Statistical Datasets, collects and republishes data from more than 75 source organizations. Data-Planet Statistical Datasets has existed since roughly 2006 and has expanded on a consistent structure during that time. For much of that time it has been independently owned; in 2018 it was purchased by SAGE Publications. The data, calculated into 16 subject areas of social and economic affairs, are principally for the period from 1950 to the present, and 70% of the data address the United States. Nevertheless, source organizations include the United Nations, IMF, World Bank, OECD, European Commission, Statistics Canada, and financial data on London, Japan, and China.

This is a subscription site, and presumably survives especially as employees at the organizations from which data is purchased are able to benefit from subscription and use its resources in carrying out statistical analysis of value to the organizations. Nonetheless, the broad availability of data and the granular level of the available data and comparisons make this product of interest to academic researchers. The FAQ tab on the Data-Planet website enables users to



Articles in vol. 21(2) and later of this journal are licensed under a Creative Commons Attribution 4.0 United States License.



This journal is published by the [University Library System, University of Pittsburgh](#) as part of its [D-Scribe Digital Publishing Program](#) and is cosponsored by the [University of Pittsburgh Press](#).

submit an inquiry about subscription rates. Website materials include a sheet comparing the attributes of Data-Planet to those of its major competitors, Statista and ProQuest.

The site has three principal sets of applications. Statistical Datasets, available since 2016, is the most versatile application. It provides multi-functional access to multiple datasets, preparing existing and calculated variables so that they can be exported to and analyzed in standard statistical packages and to GIS. Statistical Ready Reference, available since 2006, provides simpler searches and quicker results. It prepares Statistical DataSheets based on data requested by users; the DataSheets can be exported to spread sheets. Hosting Services offers Data-Planet as a host for datasets submitted by users, archiving them and integrating them into the format enabling comparison with other materials in Data Planet.

The look and feel of the Statistical Datasets interface is the greatest strength of this collection of datasets, especially because of the consistency with which data from various sources are represented. The YouTube introduction to the site by Richard Landry (<https://www.youtube.com/watch?v=J5WjSdfULTY>) provides a coherent and comprehensive introduction to the site, showing the remarkable ability to combine data from different sources and calculate new variables from them. On a single screen, one finds the subject areas and databases (on the left); the functionality for selecting and modifying individual or multiple variables (at the top); the display of data as graph, map, ranking, or pie chart (in the center); and the option of showing the data in tabular form (at the right). The tools for defining, revising, and deepening searches all appear to be available on the main screen. Among the innovative aspects of the Statistical Datasets is the “What’s Happening Here?” application. In it, for a point on a graph or chart that is an outlier or otherwise interesting, one can double-click on the point and define a Yahoo search that locates sites linked to the time, place and nature of the data.

The history of this group appears to explain much of its strength. Because the datasets come mainly from proprietary and governmental sources, they are relatively well documented and thus easier to combine. In that sense, Data-Planet is comparable to IPUMS (<https://www.ipums.org/>), which “collects, preserves and harmonizes” census microdata for the United States and most other countries. In contrast, repositories collecting and combining datasets created by academic research groups encounter the difficulty of achieving consistency among datasets created by small and often underfunded groups of researchers. Thus, ICPSR, Dataverse Network, Seshat, and the World-Historical Dataverse, while they have made some progress in collecting data, have advanced little in harmonizing data.

Richard Landry stands out as the leading figure in Data-Planet. With a BS degree in electrical engineering from Valparaiso University, he founded Conquest Systems in Bethesda, Maryland, in 1989, apparently working in preparing data for consulting work. He was joined by two co-workers

who have remained with him: in 1996 by Kari Hemdal (M.Sc. in computer science at Helsinki University), and in 1998 by Steve Braland (BS in mechanical engineering at University of Maryland). They have continued as software developers and in management roles. In 2006, Conquest Systems sold off its consulting activities and created Data-Planet. The initial work of Data-Planet was acquiring and reformatting data, and in developing DataSheets that presented graphs, maps, ranking, and pie charts. ProQuest acquired Data-Planet in 2010, but in 2012 that agreement was ended and Conquest Systems again became independent. By 2013, Data-Planet announced that it held 3.9 billion time series collections of data, and in 2014 it announced that 150,000 DataSheets had been made available.

Development of more elaborate software and acquisition of more data enabled Data-Planet to release Data-Planet Statistical Datasets as of 2016. Data-Planet announced in 2016 that it held 35 billion data points, described according to 37 metadata fields. The available YouTube tutorials are based on this expanded interface. In 2018, just as the company was acquired by SAGE Publications, it announced that it held 52 billion data points. For additional fees, Data-Planet provides access to data on market planning within the U.S., stock prices worldwide, and subnational data on China.

The Data-Planet group appears to have built patiently. They may have made good decisions at the start; they appear to have minimized changes in format along the way. The data focus heavily on U.S. national and local data, but the methodology and the interface seem ready for expansion of international coverage: the associated “libguides” include guides to statistics for 162 nations. The transformation and harmonization of diverse data appears to be a major achievement of Data-Planet; such harmonization is more difficult at cross-national levels, but the existing techniques should remain valuable. It is valuable as a way to gather quick data on social-science issues in recent times for well-documented countries. Beyond that, it is valuable as a model for dataset design and for emulation by scholars who are designing datasets—or are considering submitting datasets to Data-Planet or another repository.

About the Author: Patrick Manning is Andrew W. Mellon Professor of World History, Emeritus, at the University of Pittsburgh. He served from 2008 to 2015 as founding director of the World History Center.

Disclosure Statement: Any conflicts of interest are reported in the acknowledgments section of the article’s text. Otherwise, authors have indicated that they have no conflict of interests upon submission of the article to the journal.