What Are Our Customers Reading?
Analyzing Circulation Data to Identify Popular Subjects in a Scientific Book Collection

NIST National Institute of Standards and Technology
Technology Administration, U.S. Department of Commerce

Harriet Hassler, NIST Research Library
Information Services Division, Technology Services,
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY

INTRODUCTION

The NIST Research Library supports the work of NIST's seven laboratories specializing in Physics, Chemical Science, Materials Science, Electrical Engineering, Manufacturing Engineering, Building and Fire Research, and Information Technology. Primary users are the 1,500 scientists on staff.

The Study: Circulation data was analyzed to build a list of the *Top 100 Call Numbers* -- the most popular subjects in our book collection. A separate analysis of newer books was used to identify emerging topics of interest on the list of *Top 50 Call Numbers for New Books*.

The Purpose: The original purpose was to help focus the Library's approval plan with our book vendor. The Library was receiving about twice as many books via automatic delivery on the approval plan as it was able to purchase. The Library wanted to identify those subjects of greatest interest to our customers and only receive via automatic delivery books in very high use subject areas.

The study provided the Library with greater insight into our customers. Circulation patterns revealed the types of resources scientists used most often in the book collection, which will help the Library select materials in the future that are likely to yield a high return on investment.

SOURCE DATA REPORT

The integrated library system provided the source data for the analysis. The ILS report generated a list of the 12,594 volumes in the book collection that were last used during the two previous years. Both checkouts and in-house uses were included. Serials were excluded from the report from the outset.

Call Num.	Title	Last Used	Check-outs	In-house Uses
QA276 .B317 1989	Asymptotic techniques for statistics	3/12/2002	8	5
QA276 .B336 1984	Outliers in statistical data	9/25/2002	23	2
QA276 .T33 1990	Statistical techniques for data analysis	7/8/2003	31	10
QA276 .W28 1947	Sequential analysis	1/14/2004	30	0
R856 .B477 1981	Biomechanics of medical devices	8/30/2002	14	6
R856 .F76 1998	Frontiers in tissue engineering	2/4/2004	47	4
TK7874 .M43 2001	Microsystem technology	2/3/2004	5	0

Sections from the base ILS report

METHODOLOGY

The source data report was sorted by call number and divided by Library of Congress Classification into individual spreadsheet for each LCC class. Classes, Subclasses, and individual subjects represented by call numbers were scored.

Subclass	Call Number	Trans-actions
QA-Math.	**QA276** .B317 1989	13
QA-Math.	**QA276** .B336 1984	25
QA-Math.	**QA276** .B395 1994	16
QA-Math.	**QA276** .V22 2000	10
QA-Math.	**QA276** .T33 1990	41
QA-Math.	**QA276** .W28 1947	30

Each call number with 24 or more volumes was given a score:
Score = Transactions x Volumes / 1000

	Volumes	Transactions	Score
QA276 Mathematical Statistics, General	**85**	**1846**	**156.9**

Transactions are the total of checkouts plus in-house uses. Volumes are the number of volumes circulated in that call number. This formula was devised to create a single number that gave weight to both total circulation transactions and the total number of volumes circulated.

Cutter Numbers: Where call numbers used Cutter numbers to divide subjects into subtopics, if the Cutter number qualified for scoring independently (i.e., had 24 or more volumes), the Cutter was scored separately. If the call number had Cutter numbers but none qualified for scoring independently, then the call number was scored as a whole without regard to Cutter numbers.

THE TOP 100 CALL NUMBERS

Once all call numbers were scored, they were easily ranked.

Rank	Score	Subject Area	Topic	Call No.
1	550.3	Analytical Chemistry	Spectrum Analysis, Special	QD96
2	344.6	Optics. Light	Spectroscopy, Special	QC454
3	327.8	Computers and Special Computing	Computer Topics: Security, Client/Server, etc.	QA76.9
4	281.1	Atomic Physics	Solids. Solid State Physics	QC176.8
5	172.1	Programming	C Programming Languages	QA76.73 .C15
6	156.9	Mathematical Statistics	Mathematical Statistics, General	QA276
7	109.9	Probabilities	Probabilities, General	QA273
8	109.7	Nuclear Physics	Special Particles, Antiparticles, Families of Particles	QC793.5
9	104.8	Materials of Engineering	Polymers	TA455 .P58
10	103.1	Inorganic Chemistry	Special Inorganic Elements	QD181

The top 10 of the *Top 100 Call Numbers*

EMERGING HOT TOPICS

A separate analysis was done on the subset of books published in 1999 or later. This was intended to identify topics that were of particular high interest to our users recently. A separate list of the *Top 50 Call Number for New Books (1999-2004)* was developed. Among topics that emerged as popular for newer books were:

Nanostructure Materials • Nanotechnology • Bayesian Statistics • Microelectromechanical Systems (MEMS) • TCP/IP • Active Server Pages • Biomedical Materials • Object-Oriented Computing • Quantum Computing • Database Design

Almost all emerging topics of interest directly related to NIST Strategic Focus Areas of Nanotechnology, Biosystems and Health, and Information/Knowledge Management. These Strategic Focus Areas were adopted in 2001.

MOST CIRCULATED TITLES

From the source data report, the Library produced a list of the *Top 30 Most Circulated Titles*. This revealed that many of our most popular titles were textbooks, including some older classics. This observation fit with the results of our customer survey conducted in 2001 in which textbooks scored in the "very high use" category. Computer and IT manuals were also popular.

The most circulated book in the collection was *The Waite Group's C++ Programming*, 2nd ed, by John Thomas Berry. The second most circulated was Draper & Smiths *Applied Regression Analysis*.

CONCLUSION

This report suggests that NIST scientists use the book collection to learn about topics that support their research, such as programming, research methodology, and statistical analysis, as much, if not more than to learn about their specific area of expertise. Studies have shown that journals, not books, are the single most important information resource for maintaining current awareness for scientists and engineers. At NIST, the book collection is more likely to be used to investigate fields in which the scientist is less familiar, but in which a situational need for more knowledge has occurred. This thesis is supported by the popularity of textbooks, which provide general overviews of a subject.

The Library will use the results of this study to direct the limited amount of resources available to purchase titles that are likely to be highly used. The Library is also planning to purchase new editions of popular, heavily used textbooks to keep the collection current.