

PREDICTING THE NEED FOR MULTIPLE COPIES OF BOOKS

Robert S. GRANT, Presently at Hope College, Holland, Michigan.

*An industrial inventory technique adapted to a university library's computer based circulation system as one aid in identifying heavily used books for multiple-copy purchase.*

The University of Windsor has approximately 5,000 students. The University Library's open stacks contain more than 300,000 volumes, 100,000 of which are non-circulating (bound periodicals and reference books). There are approximately 200,000 books available for circulation, a books-to-student ratio of 40:1. Nevertheless, a perennial student complaint is: "Why is it that every time I need a book, someone else has already checked it out?"

To help mitigate this problem, the Library decided several years ago to embark upon a programme of purchasing multiple copies of much used books. The question then became one of determining which books would need duplicating, and how many more copies of each title would need to be bought.

Suggestions of titles to be duplicated were at first solicited from the faculty, but ever-increasing demands on them prevented their being more than minimally cooperative.

Three years ago, in an effort to increase the availability of books to undergraduates, the Library changed its circulation period for undergraduates from two weeks to one week, with unlimited renewals. At the same time there was instituted a system whereby a student filled out a reserve card requesting that he be allowed to check out a book upon its

return. When there were five or more such requests, then a copy of the book was to be purchased. Although this system of ordering multiple copies was very cumbersome, it was better than nothing.

An article by William L. Leffler (1) suggested a system of adapting industrial inventory techniques to the problem of identifying books to be duplicated that would be compatible with the Library's computer based circulation system and also could be expected to be simpler and more thorough than the above method of buying multiple copies.

Without rehearsing Leffler's arguments, the basic formula used in this project can be simply stated as:

$$N_{\text{Books}} = \frac{N \times N_{95\%}}{T}$$

where  $N_{\text{Books}}$  = the number of copies of a single title necessary to meet at least 95% of student demand for that title;  $T$  = number of days of observation, i.e., the number of days in the academic year in which students are permitted to check out books (a constant of 273 in this formula, being the number of days in the period from 1 September to 31 May);  $N$  = total number of times a title circulated during  $T$ ;  $N_{95\%} = \alpha + 2S$ , where  $\alpha$  = the average length of time a title was on loan, i.e., the total number of days in which a title was in circulation divided by the number of times ( $N$ ) the title circulated.  $S$  = Standard Deviation, which is computed as:

$$\sqrt{\frac{\sum (A_1 - \alpha)^2}{N}}$$

$A_1$  is the length of time, in terms of days, that a single title was off the shelves each time it circulated, and is not to be confused with  $\alpha$  which is the average length of time (over the academic year) that the same title was on loan.

The sum ( $\sum$ ) of all the  $A_1$ 's was used earlier to calculate  $\alpha$ .

$$\alpha = \frac{A_1 + A_2 + A_3 \dots \text{etc.}}{N}$$

For example, if a book circulated three times during the academic year (the first time for 18 days, the second time for 20 days, and the third time for 3 days) then  $\alpha$  (the average length of time the book was on loan) would be calculated as

$$\frac{18 + 20 + 3}{3}, \text{ or } 13.66$$

At this point it should be noted that although the Library continues to accept request cards for books presently on loan (and to reserve books for the requestors), these requests are not used as part of the data in determining the number of copies necessary to meet at least 95% of the demand. For one thing, there is no way of knowing how long the person making a request will want to keep a book out, and time is an important element in the formula. But more importantly, the formula, as it now stands, attempts to account for unsatisfied requests. It assumes that in at least some instances there will be more requests for a title than there

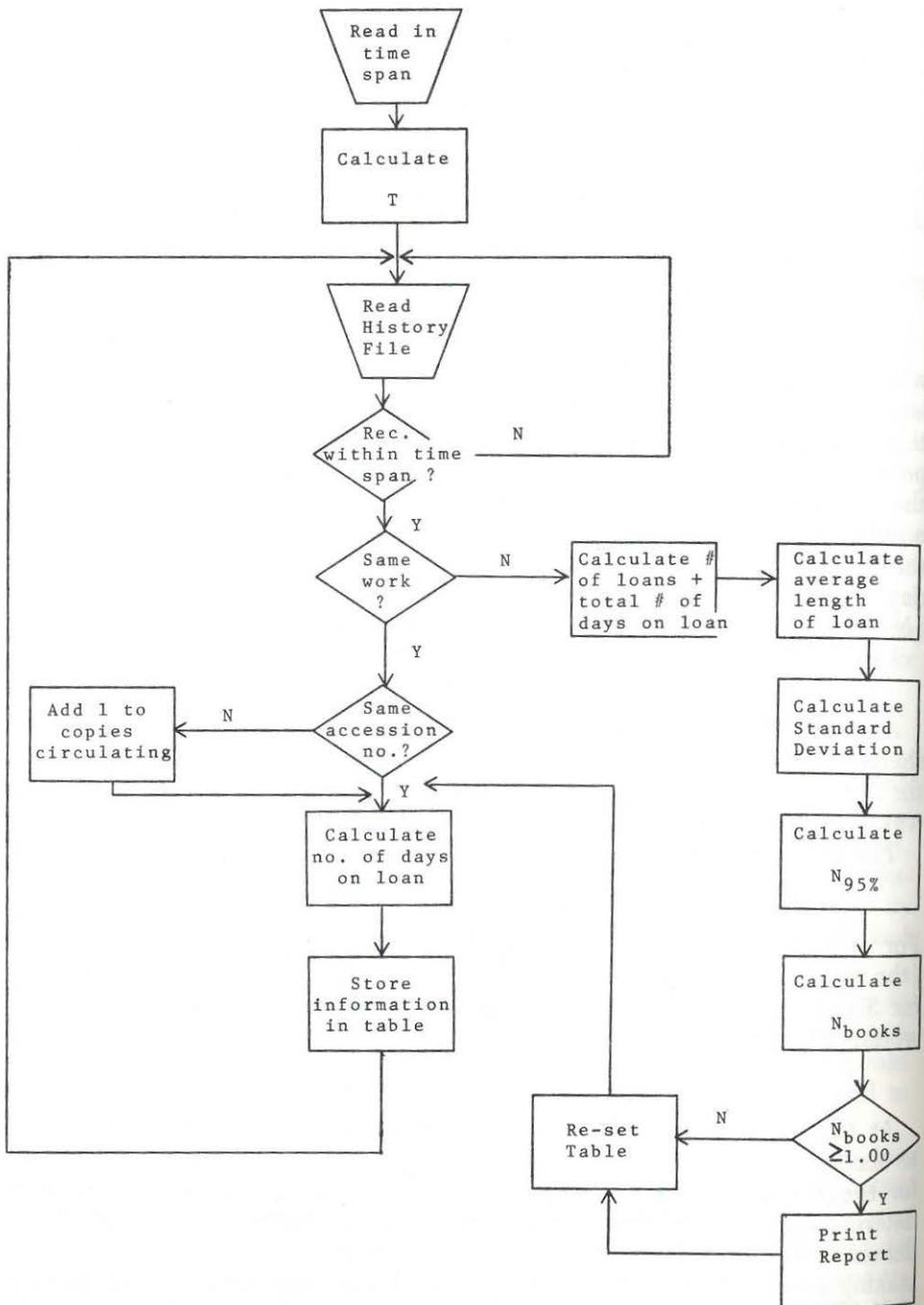


Fig. 1. Programme Logic.

are copies in the Library. By providing an analysis of the present circulation profile of each book, the formula attempts to predict the number of copies of each title the Library would need to have in order to more adequately accommodate unsatisfied demand.

The programme for performing the calculations is written in PL/I and is run on an IBM 360/50 (Figure 1). The execution time for 140,000 circulation records (each time a book circulates the data on its circulation is considered a single record) is 15 minutes.

The Historical Record File, the source of data for the programme, is incremented each time a book in circulation is returned. Figure 2 shows the format of this file.

The file itself is a sequential file stored on magnetic tape, updated daily to include the previous day's circulation data. Entries are arranged in LC call number-accession number order.

FIELD	LENGTH	ACCUMULATIVE LENGTH
Card Type	1	1
LC Call Number	29	30
Author	15	45
Accession Number	6	51
Spare	1	52
Card Sequence Number	6	58
Spare	2	60
Borrower's ID Code	1	61
Borrower's ID Number	6	67
Spare	3	70
Action Code	1	71
Due Date (MMDDYY) (MO.-DAY-YR.)	6	77
Spare	3	80
Indicator	1	81
Date Charged Out (YYDDD) (YR.-DAY)	5	86
Date Returned (YYDDD) (YR.-DAY)	5	91

Fig. 2. Format of Historical Record File.

## RESULTS

After the calculations described above have been performed for every title circulated during the academic year, a print-out of the results is produced (Figure 3). In order to limit paperwork, only those results under "Projected Need" which were  $\geq 1.00$  appear on the print-out; any results less than 1.00 were suppressed.

The column labelled "Transactions" is simply the number of times the book was checked out and checked back in again. The column, "Average Loan Period" is the  $\alpha$  described in the formula above. And the column, "Copies Circulated" is the number of books with the same classification number as listed on the left-hand column, but with different accession numbers, checked out during the year. This figure is not the number of copies of the book that the Library owns, which could, in some instances, be more copies than were actually circulated.

The column labelled "Projected Need" should, according to the calculations, indicate the number of copies of a title which could accommodate the demand for that title with 95% certainty. In order to find out whether or not the Library should purchase more copies of a particular title, the number listed in this column is simply checked against the number of

CLASSIFICATION	AUTHOR	PROJECTED NEED	TRANS.	AVG LOAN PERIOD	COPIES CIRCUL.
AM---101...C34B8--	CANADA-NATIONAL	3.61	37	10.45	17
B----56...C6-----	COLLINS-JAMES-D	1.14	21	8.00	2
B----65...B6-----	BODENHEIMER, E.-	1.21	12	11.50	3
B----65...R6-----	ROMMEN-HEINRICH	1.34	5	20.60	2
B----67...B58-----	BLAKE-RALPH-M--	2.00	4	36.75	2
B----67...N22-----	NAGEL-ERNEST---	2.34	23	11.39	3
B----72...C63-----	COPELSTON-F.C.-	2.39	27	9.18	10
B----72...H5-----	GILSON-E.H.----	1.64	26	9.03	14
B----72...J6-----	JOAD-CYRIL-EDWI	2.84	8	21.75	2
B----72...P3-----	PARKER-F.H.----	2.48	4	41.00	2
B----358...C57-----	PLATO-----	5.68	21	15.61	3
B----358...J8-----	PLATO-----	2.00	38	8.07	10
B----358...W7-----	PLATO-----	2.72	5	39.80	3
B----377...A2B5---	PLATO-----	3.65	8	35.37	2
B----378...A2C6---	PLATO-----	1.58	2	73.00	2
B----381...A5T35---	PLATO-----	1.04	3	36.33	3
B----385...A6-----	ANDERSON-F-H---	2.92	16	13.43	2
B----395...B77-----	BRUMBAUGH-R-S--	2.05	12	13.33	1
B----395...C6-----	CROMBIE-I-----	3.02	17	12.41	2
B----395...C67-----	GRUBE-GEORGE-M-	5.13	30	10.30	5
B----395...G78-----	GUARDINI, R.----	2.04	17	12.23	4
B----395...K6-----	KOYRE-ALEXANDRE	1.13	4	21.75	1
B----395...L6-----	LODGE-RUPERT-C-	1.88	3	51.33	1
B----395...S53-----	SHOREY, PAUL---	4.69	23	11.91	4
B----398...T25-----	TAYLOR, A.E.----	1.31	28	7.75	5
B----398...E8H17--	HALL, ROBERT-W.-	2.99	11	16.72	2
B----407...L8L9---	LUTOSLAWSKI, W.-	3.10	4	59.25	1
B----505...M2-----	ARISTOTELES, ---	2.88	17	12.00	7
B----505...O3-----	OATES-W.J.-----	3.86	9	27.33	7
B----528...Z413---	ZELLER-EDUARD--	1.39	6	33.00	2
B----528...P751---	POHLENZ-MAX----	1.35	5	34.60	2
B----667...S25-----	SAMBURSKY-SAM--	1.36	5	42.40	1
B----701...D4D6---	DONDAINE, H.F.--	1.03	2	69.00	1
B----701...A4E5---	PROCLUS-DIADOCH	1.11	2	72.50	1

Fig. 3. Circulation History Analysis Report.

copies listed for this classification number in the official shelf list. For example, the book classified as B.72.J6 shows a "Projected Need" of 2.84. Therefore if the Library had three copies of this book, and the book's circulation pattern did not change significantly in the immediate future, then the Library would be able to fill 95% of the requests. The official shelf list, however, indicates that the Library only owns two copies of this title, suggesting that at least one more copy should be purchased to meet present demand. These calculations do not anticipate future demand on the book. Also, doubling the number of copies can never succeed in doubling circulation, a fact demonstrated by Leimkuhler (2). This print-out, therefore, can only serve as one guide to multiple-copy purchase.

### PRECAUTIONS AND PITFALLS

In using the results of these computations as a guide to the purchase of multiple copies, the Librarian should be aware of several factors which may have distorted the results. One is that the student who checks out the only copy of a book and keeps checking it out all year, in lieu of buying his own copy, creates a false "demand" for the book. It may be that he is the only person in the University interested in it, and when he graduates this book may sit out its life on the shelves completely unused.

However, since the Historical Record File contains the Borrower's ID number, it is possible to distinguish between an original loan and a renewal. The first time the Borrower's ID number appears on the book's circulation record indicates the original loan. Each additional and consecutive time the same Borrower's ID number appears on the same circulation record indicates a renewal. Although the pilot project did not contain provisions for obviating this problem, it would have been simple enough to build into the programme a mechanism for suppressing the unwanted data.

A faculty member who assigns parts of books for students to read, but does not place the books on reserve, forces competition for them on the open shelves. This too creates a demand which may not exist after the professor leaves the University or stops teaching a particular course. The librarian should be aware of such possible short-lived demands that may never recur.

The circulation analysis programme was executed at the end of one academic year in order to provide the University of Windsor librarians with guidelines for purchase of multiple copies of books to be used in the next academic year. If it were known that a particular book receiving heavy use one year would not receive equally heavy use in the next (because, for example, the particular course requiring that book would no longer be taught; or the book would be placed on a "two-hour reserve" for the coming academic year; or the book circulated frequently in one year only because it was on the "best-seller list"), then it would be folly to purchase three or four additional copies of the book just because the computer print-out indicated that a number of additional copies were

needed. Other factors, therefore, although not included in the input data, are certainly relevant in determining the need for multiple copies.

At the University of Windsor Library, a book that needs to be re-bound because of heavy use or mutilation is charged out to the Bindery Department. It then shows up on the Historical Record File, just as though it had been charged out. But since the "Borrower's ID number" for books charged to the Bindery Department consists of all zeroes, it would be simple enough to identify and suppress these particular records as unwanted data.

### BY-PRODUCTS

In addition to providing a list of books to be considered for duplication, the Historical Record File upon analysis revealed several other interesting facts about the University Library's circulation.

Most noteworthy is the fact that, although there were more than 200,000 circulating books sitting on the open shelves at the time of this pilot project, only 40,205 different titles circulated for a total of 134,276 times. Assuming there were only 100,000 different titles among the 200,000 books, this would mean that nearly 60% of the collection was probably not used by the students.

Of the 40,205 different works which did circulate, the calculations indicated that only 3,257 titles required one or more copies in order to fill 95% of the requests. Of this latter number, only 570 titles were in need of duplication. (That is to say, the number of copies listed under projected need exceeded the number of copies actually owned by the Library as indicated by the shelf list.)

A random sample comprising one-third of these 570 titles was checked to see whether or not the books were in print. Indications were that 38% of the titles in need of duplication were no longer in print.

### CONCLUSIONS

A close examination of the 570 titles apparently in need of duplication reveals that, with very few exceptions, students are apparently checking out only books that are curriculum oriented in the most narrow sense, i.e., books which they need to use in writing term papers. Nevertheless, one can appreciate the fact that these books are in demand by the student, and if the Library is to be responsive to users' demands on its facilities, it will need to spend part of the book budget each year purchasing multiple copies of the most heavily used books.

Unfortunately, even with these good intentions and the sophisticated assistance of the computer, students' demands for books will still be frustrated (at least one-out-of-three times) because books which need to be duplicated are no longer in print.

### PROGRAMME

A print-out copy of the circulation analysis programme described above

is available from Mrs. Jean Griffiths, Computer Centre, University of Windsor, Windsor, Ontario, Canada.

#### ACKNOWLEDGMENTS

The initial impetus and continuous guidance for this project was provided by Albert V. Mate, Assistant Librarian for Public Services at the University of Windsor. Dr. Martin Basic, Faculty of Business Administration, acted as consultant. Systems analyst was Mrs. Jean Griffiths, and programmer was Mrs. Lillian Jin, both at the University Computer Centre.

#### REFERENCES

1. Leffler, William L.: "A Statistical Method for Circulation Analysis," *College and Research Libraries*, 25 (1964), 488-490.
2. Leimkuhler, Ferdinand F.: "Systems Analysis in University Libraries," *College and Research Libraries*, 27 (1966), 13-18.