Analysis of data collected from 3,000 academic libraries by the 1977 Library General Information Surveys reveals that library operating budgets, institutional enrollment, and library circulation are the best predictors of reference and directional transactions. Fifty-five percent of the transactions at reference service points are directional, while the remainder are reference transactions; university libraries report significantly higher numbers of transactions than either 4-year or 2-year colleges; and publicly controlled colleges report greater numbers than private institutions. Similarly, reference and directional transactions vary with total operating budget, collection and staff size, and enrollment. The picture is complicated, however, by intervariable relationships; e.g., university libraries tend to have larger operating budgets, staff, and collections. Under these conditions, regression analysis is a better procedure to predict the number of reference and directional transactions. (Author/RAA)
ANALYSIS OF REFERENCE STATISTICS REPORTED
IN 1977 LIBRARY GENERAL INFORMATION
SURVEY

Neal K. Kaske
OCLC, Inc.

and

Rao Aluri
SUNY/Buffalo

June 1980

Paper Presented to Library Research Round Table,
Annual Conference of the American Library Association,
New York City, June 29, 1980

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY
Rao Aluri"

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)"
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>LIBGIS (Library General Information Survey) Data</td>
<td>3</td>
</tr>
<tr>
<td>Objective of the Study Data</td>
<td>5</td>
</tr>
<tr>
<td>Data</td>
<td>6</td>
</tr>
<tr>
<td>Analysis</td>
<td></td>
</tr>
<tr>
<td>A. Preliminary Analysis</td>
<td>7</td>
</tr>
<tr>
<td>B. Regression Equations</td>
<td>10</td>
</tr>
<tr>
<td>a. Correlation between Variables</td>
<td>12</td>
</tr>
<tr>
<td>b. Regression Analysis</td>
<td>14</td>
</tr>
<tr>
<td>i. Reference Transactions</td>
<td>16</td>
</tr>
<tr>
<td>ii. Directional Transactions</td>
<td>18</td>
</tr>
<tr>
<td>iii. All Transactions</td>
<td>20</td>
</tr>
<tr>
<td>iv. Regression Forced through zero</td>
<td>22</td>
</tr>
<tr>
<td>Conclusions</td>
<td>24</td>
</tr>
<tr>
<td>References</td>
<td>27</td>
</tr>
</tbody>
</table>
Tables and Figures

Table 1. Transactions by Type of Control ..... 8
Table 2. Transactions by Type of Institution ..... 8
Table 3. Transactions by Level of Enrollment ..... 9
Table 4. Pearson Correlation Coefficients ..... 13
Table 5. Partial Correlation Coefficients ..... 15

Figure 1. Plot of Reference Transactions vs.
Library Operating Expenditures ... 19
Figure 2. Plot of Directional Transactions vs.
Library Operating Expenditures ... 21
Figure 3. Plot of All Types of Transactions vs.
Library Operating Expenditures ... 23
I. Introduction:

The need for measurement and evaluation of reference service for such purposes as budget justification, determination of reference staffing patterns, and evaluation of reference collection has been recognized for a long time. Until recently, however, two problems proved to be stumbling blocks in making advancement in this area. These are: (1) absence of standardization for measuring reference activity; and (2) lack of meaningful sets of standards or nation-wide norms for reference service against which individual libraries could compare their performance.

Libraries differ considerably in the type of information collected to measure the reference service. While some libraries keep a simple count of the questions asked at the reference desks, other libraries developed elaborate schemes to categorize the reference questions. Examples of these categorizations include the type of reference question (e.g., ready reference, directional, and research); source of question (e.g., telephone call); and subject area (e.g., literature, social sciences and physical sciences). The disadvantage of these locally designed schemes for measuring reference service is that the information obtained at one library is not transferable to other libraries. Libraries, for example, may widely differ in interpreting a given question as a ready reference or research question. Similarly, the data collected in a general library of
a four-year college as to the distribution of reference questions by subject area is not applicable to a library serving a different clientele, say, engineering students. Complicating this issue further is the fact that there is not uniformity among libraries as to who provides reference service or what tasks are performed by reference departments and, conceivably, what activity should be reported as reference transaction. 3

Since there was never an agreement on what to measure and how to measure, no serious attempt was ever made to develop nation-wide norms or standards for reference activity. Rothstein points out, "no area or library service has been more deficient in such standards than reference service." 4 The disadvantage of the lack of reference standards is the inability to assess, however crudely, the effectiveness of reference service at a given library. For example, the information that Library A reports 250 reference transactions per day does not mean much to Library B which reports, say, 50 transactions per day. Such a difference in the absolute number of transactions does not automatically mean that the performance of Library B is worse than that of Library A. In fact, this difference could have arisen because of differing library sizes and clientele. In other words, the information on the number of reference transactions at Library B is just not enough to evaluate its performance in relation to that of Library A.
II. LIBGIS (Library General Information Survey) Data:

The task of standardizing the measurement of reference service was, finally, tackled by the ALA Committee on Statistics for Reference Service which is, currently, a part of the Statistics Section of the ALA's Library Administration and Management Association (LAMA). On the recommendation of this Committee, National Center for Education Statistics (NCES) started incorporating two questions on reference transactions beginning 1976. These two questions, respectively, seek to collect the data on the number of reference transactions per typical week and the number of directional transactions per typical week. The 1977 LIBGIS questionnaire provides the following instructions and definitions regarding the reference and directional transactions.

REFERENCE AND DIRECTIONAL TRANSACTIONS. Report contacts of all main branch library personnel whose assigned duties include provision of reference/information service. A staff member should report each contact separately, whether or not the user has already consulted either that staff member or another on the same information need. A contact which includes both reference and directional service is one reference transaction. Duration should not be an element in determining whether a transaction is reference or directional.
TOTAL REFERENCE TRANSACTIONS PER TYPICAL WEEK. A Reference transaction is an information contact which involves the use, recommendation, interpretation, or instruction in the use of one or more information sources, or knowledge of such sources, by a member of the reference/information staff. Information sources include:

1. print and nonprint materials;
2. machine readable data bases (including computer assisted instruction);
3. library bibliographic records, excluding circulation records;
4. other libraries and institutions; and
5. persons both inside and outside the library.

A question answered through the utilization of information gained from previous consultation of such sources is considered a reference transaction even if the source is not consulted again.

TOTAL DIRECTIONAL TRANSACTIONS PER TYPICAL WEEK. A directional transaction is an information contact which facilitates the use of the library in which the contact occurs, and its environs, and which may involve the use of sources describing that library, such as schedules, floor plans, handbooks, and policy statements. Examples of directional transactions are:

1. directions for locating facilities such as restrooms, carrels, and telephones;
2. directions for locating library staff and users;
3. directions for locating materials for which the user has a call number;
4. supplying materials such as paper and pencils; and
5. assisting users with the operation of machines.

The significance of this simple categorization of informational contacts on the one hand and the provision of detailed instructions and examples of categorizations on the other is that (1) it provides the first nation-wide standardization of these categories which represent a broad consensus; and (2) the problem of widely differing interpretations is lessened. This standardization, of course, is bought at a price. The LIBGIS categorization of informational contacts into reference and directional transactions is primitive. The reference transactions, one of the LIBGIS categories, does not
distinguish between a question which can be answered readily by using standard reference sources such as almanacs and dictionaries and a question that needs consulting a multitude of sources requiring hours of work. Similarly, the LIBGIS data does not attempt to gauge the quality of information and service provided by reference personnel. Despite these drawbacks, LIBGIS questions on reference and directional transactions represent a significant advancement for the measurement of reference service.

LIBGIS also represents an advancement in the area of evaluation of reference service. For the first time, library community has access to reference data, which has been collected under generally similar conditions following similar categorizations. Coupled with this data, LIBGIS also provides data on other aspects of libraries such as their operating budgets, staff sizes and sizes of their clientele. All this quantitative information can now be used to determine the quantitative norms for reference service. These norms, in turn, can be used to evaluate the level of service at individual libraries.

III. Objective of the Study:

The objective of the present study is to examine the relationships between reference and directional transactions
on the one hand and the variables such as library operating budgets, collection and staff sizes on the other with a view of fitting this data into a regression equation. This regression equation, once developed, would assist the individual libraries to calculate the expected number of reference and directional transactions in their institutions and to compare these expected values with the actual number of reference and directional transactions at their reference desks.

IV. Data:

The data for this project has been derived from that reported by the 1977 Library General Information Survey (LIBGIS). The 1977 LIBGIS reports data on 3,149 academic institutions and the data is recorded on a 9-track, 1600 BPI, odd parity, EBCDIC tape. Since the original data is far more detailed than is suitable or necessary for this project, part of it has been extracted and transferred to a separate file. This subset of data gives, for each institution, college enrollment, total operating expenditures, number of items circulated, reference and directional transactions per typical week, hours the library is open per typical week, titles held and added, periodicals held and added, professional and total staff sizes. Also included are indicators showing whether the library is a part of publicly or privately controlled institution and whether it is a university, four-year or two-year college library. Information on this file is analyzed using Statistical Package for Social Sciences (SPSS).
V. Analysis:

A. Preliminary Analysis:

The average number of reference transactions per typical week (hereafter, reference transactions) reported by 1977 LIBGIS is 315.8 (mode=50.0, median=120.1, range=7,366, and 95% C.I.=291.2 to 340.4). The average number of directional transactions per typical week (hereafter, directional transactions) is 398.2 (mode=100.0, median=145.2, range=8,830, 95% C.I.=365.8 to 430.6). These values are obtained by treating those institutions which reported either zero or 10,000 or more transactions per week as missing cases. There are 480 and 564 missing cases respectively in the calculation of the averages of reference and directional transactions all but a few of which reported zero transactions. As can be seen from these averages, directional transactions slightly outnumber the reference transactions.

As an initial step, one can show that the number of reference and directional transactions are dependent on the institutional characteristics such as the type of control (i.e., public or private), type of institution (i.e., university, four-year or two-year), and the level of enrollment. Table 1 shows that public institutions reported larger number of reference and directional transactions than private institutions. Number of transactions reported by universities, as can be seen from Table 2, are much higher than those reported by four-year and two-year colleges. Similarly, Table 3 demonstrates that the number of transactions increase with the level of enrollment.
Table 1
Transactions by Type of Control

<table>
<thead>
<tr>
<th>Type of Control</th>
<th>Average Number of Reference Transactions</th>
<th>N</th>
<th>Directional Transactions</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public</td>
<td>429.9</td>
<td>1316</td>
<td>521.0</td>
<td>1275</td>
</tr>
<tr>
<td>Private</td>
<td>204.8</td>
<td>1353</td>
<td>278.6</td>
<td>1310</td>
</tr>
</tbody>
</table>

F = 82.95  
\(\text{df} = 1, 2667\)  
\(p < 0.0000\)

F = 54.82  
\(\text{df} = 1, 2583\)  
\(p < 0.0000\)

Table 2
Transactions by Type of Institution

<table>
<thead>
<tr>
<th>Type of Institution</th>
<th>Average Number of Reference Transactions</th>
<th>N</th>
<th>Directional Transactions</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>University</td>
<td>1558.3</td>
<td>140</td>
<td>1876.8</td>
<td>140</td>
</tr>
<tr>
<td>Four-Year</td>
<td>263.7</td>
<td>1503</td>
<td>324.2</td>
<td>1452</td>
</tr>
<tr>
<td>Two-Year</td>
<td>222.5</td>
<td>1026</td>
<td>297.9</td>
<td>993</td>
</tr>
</tbody>
</table>

F = 342.76  
\(\text{df} = 2, 2666\)  
\(p < 0.0000\)

F = 278.4  
\(\text{df} = 2, 2582\)  
\(p < 0.0000\)
Table 3
Transactions by Level of Enrollment

<table>
<thead>
<tr>
<th>Reference Transactions</th>
<th>Average Number of Transactions</th>
<th>N</th>
<th>Directional Transactions</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 199</td>
<td>93.2</td>
<td>193</td>
<td>86.2</td>
<td>184</td>
</tr>
<tr>
<td>200 to 499</td>
<td>89.0</td>
<td>282</td>
<td>116.2</td>
<td>275</td>
</tr>
<tr>
<td>500 to 999</td>
<td>138.0</td>
<td>503</td>
<td>215.5</td>
<td>488</td>
</tr>
<tr>
<td>1000 to 2499</td>
<td>178.7</td>
<td>739</td>
<td>246.0</td>
<td>718</td>
</tr>
<tr>
<td>2500 to 4999</td>
<td>282.2</td>
<td>376</td>
<td>390.0</td>
<td>366</td>
</tr>
<tr>
<td>5000 to 9999</td>
<td>524.5</td>
<td>310</td>
<td>659.1</td>
<td>300</td>
</tr>
<tr>
<td>10000 to 19999</td>
<td>902.9</td>
<td>179</td>
<td>945.7</td>
<td>168</td>
</tr>
<tr>
<td>20000 and over</td>
<td>1930.6</td>
<td>87</td>
<td>2330.0</td>
<td>86</td>
</tr>
</tbody>
</table>

$F = 177.3$  
$df = 7, 2661$  
$p < 0$

$F = 125.5$  
$df = 7, 2577$  
$p < 0$
While this information is interesting, it does not help an individual library to arrive at the expected number of transactions at its reference desk. For instance, the averages of reference and directional transactions are only of minimal help because the number of transactions reported by three out of four libraries fall below them. Similarly, the other averages shown in Tables 1, 2, and 3 are not sufficient to estimate the expected number of transactions. This is because a given institution may be a two-year college with an enrollment of over 20,000. Tables 2 and 3 suggest vastly differing averages for such an institution. One way to overcome this problem is to fit the reference and directional transactions into regression equations involving a number of independent variables. Such regression equations are far more helpful in predicting the expected number of transactions for a given institution if the values of the independent variables included in the regression equations are known. The second part of this section attempts to develop such regression equations for reference and directional transactions.

B. Regression Equations for Reference and Directional Transactions:

Eight independent variables, namely, size of enrollment, total operating expenditures, circulation, hours open per typical week, titles added, titles held, professional
staff size and total staff size, are targeted as potentially useful predictors of reference and directional transactions. Size of enrollment is an indicator of the size of the library clientele and it is reasonable to assume that a larger clientele would give rise to larger number of transactions at reference service points. Operating expenditures indicate monies allotted for personnel and library materials. Again, larger operating budgets might indicate greater ability and willingness on the part of the libraries to support larger reference staffs which, in turn, may encourage larger number of transactions. Similar arguments can be extended to support the inclusion of other variables.

The following paragraphs describe each of these independent variables.

Size of enrollment: It is the total number of students enrolled for the survey year, 1977. This data is not obtained from LIBGIS questionnaire but is taken from the National Center for Education Statistics' Institutional Characteristics File.

Total operating budget (or expenditures): It is the sum of library's expenditures for salaries and wages, materials, binding and rebinding, and travel and other miscellaneous expenses.

Circulation: It is the number of transactions of all materials charged out to library users. Excludes dial access transactions interlibrary loans.

Hours open per typical week: LIBGIS questionnaire defines it as "the total number of hours the central or main library is open for general users in a typical week." Typical week is "a week containing no holidays, during a regular semester i.e., fall or spring, in which the central or main library is open its regular hours for general users."
Titles added during academic year and Titles held at end of academic year: For the purpose of these two variables, a title is "a publication which forms a separate bibliographic whole, whether issued in one or several volumes, reels, discs, slides, or parts. For this project, these figures include only "bookstock." Bookstock, according to LIBGIS questionnaire is the "cataloged collection(s) of books and other printed materials that are cataloged in the same manner as books." Includes bound periodicals and cataloged government documents but excludes such items as separate government documents, pamphlets, technical reports and others which "are not cataloged in the same manner as books."

Professional staff: Number of full time equivalent professional staff including staff such as Chief, Deputy, etc. librarians who have "administrative responsibilities for management of the library"; librarians who do work "that requires professional training and skill in the theoretical or scientific aspect of library work"; and other staff such as curators, archivists, media specialists, and computer specialists.

Total staff: Number of full time equivalent employees including professional staff and nonprofessional supporting staff. Excludes maintenance, custodial, and student employees.

a. Correlation between variables:

Once the two dependent and eight independent variables are identified, the next step is to examine the strength of relationship between these variables. Table 4 shows the Pearson correlation coefficients between the ten variables under consideration. Both reference and directional transactions are highly correlated with all the independent variables with the exception of hours the library is open. The variable, library operating expenditures, shows the best relationship with reference and directional transactions. Other variables such as enrollment and circulation also exhibit strong relationship with the dependent variables. However, these variables simultaneously exhibit much stronger relationship with the library operating expenditures. Consequently, the strength of
Table 4
Pearson Correlation Coefficients

(p=0.001)

<table>
<thead>
<tr>
<th>Enrollment</th>
<th>Expenditures</th>
<th>Circulation</th>
<th>Reference Transactions</th>
<th>Directional Transactions</th>
<th>Hours Open</th>
<th>Titles Added</th>
<th>Titles Held</th>
<th>Professional Staff</th>
<th>Total Staff</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>.74</td>
<td>.72</td>
<td>.58</td>
<td>.53</td>
<td>.30</td>
<td>.66</td>
<td>.61</td>
<td>.73</td>
<td>.73</td>
</tr>
<tr>
<td>1.00</td>
<td>.90</td>
<td>.63</td>
<td>.55</td>
<td>.34</td>
<td>.90</td>
<td>.92</td>
<td>.96</td>
<td>.98</td>
<td>.98</td>
</tr>
<tr>
<td>1.00</td>
<td>.61</td>
<td>.54</td>
<td>.33</td>
<td>.81</td>
<td>.83</td>
<td>.86</td>
<td>.89</td>
<td>.61</td>
<td>.61</td>
</tr>
<tr>
<td>1.00</td>
<td>.74</td>
<td>.23</td>
<td>.55</td>
<td>.54</td>
<td>.60</td>
<td>.61</td>
<td>.53</td>
<td>.53</td>
<td>.53</td>
</tr>
<tr>
<td>1.00</td>
<td>.23</td>
<td>.48</td>
<td>.47</td>
<td>.52</td>
<td>.53</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
</tr>
<tr>
<td>1.00</td>
<td>.36</td>
<td>.33</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
<td>.86</td>
<td>.88</td>
<td>.88</td>
<td>.88</td>
</tr>
<tr>
<td>1.00</td>
<td>.90</td>
<td>.86</td>
<td>.88</td>
<td>.88</td>
<td>.88</td>
<td>.92</td>
<td>.92</td>
<td>.92</td>
<td>.92</td>
</tr>
<tr>
<td>1.00</td>
<td>.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.97</td>
<td>.97</td>
</tr>
<tr>
<td>1.00</td>
<td>.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the relationship between, say, circulation and reference transactions may be the result of the relationship between circulation and library expenditures. To examine this possibility, one can compute the partial correlation coefficients between the dependent variables while controlling for the effects of the operating expenditures. Table 5 shows these partial correlation coefficients. Once the effects of library operating expenditures are controlled for, with the exception of enrollment and circulation, all the other variables show negligible and, in fact, negative relationship with the reference and directional transactions. The partial correlation coefficients, thus, show that library operating expenditures, enrollment, and circulation are the most important variables for predicting reference and directional transactions.

b. Regression analysis:

The regression analysis attempts to arrive at an equation of the form:

\[ Y' = A + B_1 X_1 + B_2 X_2 + \ldots \]

where \( Y' \), in our case, is the estimated number of reference (or, directional transactions); \( X_1, X_2 \), etc. are independent variables such library operating expenditures in dollars, student enrollment, etc.; \( B_1, B_2, \ldots \) are partial regression coefficients; and \( A \) is a constant. In the previous section on partial correlation coefficients, three variables, namely, library operating expenditures, enrollment, and circulation are identified as exhibiting relatively stronger relationship with reference and directional transactions.
Table 5
Partial Correlation Coefficients
(Controlling for the effects of Library Operating Expenditures)

<table>
<thead>
<tr>
<th>Reference Transactions</th>
<th>.23</th>
<th>.15</th>
<th>.02</th>
<th>-.02</th>
<th>-.12</th>
<th>-.03</th>
<th>-.02</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Directional Transactions</th>
<th>.22</th>
<th>.13</th>
<th>.05</th>
<th>-.03</th>
<th>-.10</th>
<th>-.02</th>
<th>-.03</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Enrollment</th>
<th>Circulation</th>
<th>Hours Open</th>
<th>Titles Added</th>
<th>Titles Held</th>
<th>Professional Staff</th>
<th>Total Staff</th>
</tr>
</thead>
</table>
Consequently, these three variables are used for developing regression equations to predict the reference and directional transactions.

i. Reference transactions:

Regression equation involving operating expenditures, enrollment, and circulation on the one hand and reference transactions on the other is:

\[
\text{Estimated number of reference transactions} = 90.1 + 0.18(\text{Operating expenditures in } \$1,000s) + 0.027(\text{number of students}) + 0.82(\text{items circulated in } 1,000s).
\] ....(1)

This equation is able to explain 43% of variation in the reference transactions (multiple correlation=0.656, \( R^2 = 0.43 \)) and the \( F \) values for all the partial regression coefficients and the constant are highly significant (\( p < 0.001 \) or 0.000). This equation tells that for every increase of $1,000 in operating expenditures, one can expect an increase of 0.18 reference transactions per typical week; for every increase of 1,000 students, one can expect an increase of 27 reference transactions; and there will be an increase of 0.82 transactions for every increase of 1,000 circulated items.

The usefulness of equation 1 can be seen from the following example. In 1977, New Jersey Institute of Technology had a student enrollment of 5,042, operating budget of $463,637, and circulation of 48,358 items. One can plug this data into
equation 1 to compute the expected number of reference transactions. The computation would yield:

\[
\text{Expected number of reference transactions} = 90.1 + 0.18 \times 464 + 0.027 \times 5042 + 0.82 \times 48
\]

\[= 349.1\]

The observed number of transactions for the New Jersey Institute of Technology in 1977 were 200 per week. The knowledge of the expected number of reference transactions and the actual number of transactions would, then, serve as a starting point for the library to evaluate its reference service.

To examine the contribution of other independent variables to the explanation of reference transactions, a step by step regression procedure was used. This analysis indicated that all other remaining variables (i.e., other than the three primary independent variables included in equation 1) together explained less than 1% variation in reference transactions. In fact, the total operating expenditures alone explained 39% of the variation (multiple correlation=0.625, \(R^2=0.39\)) in reference transactions while enrollment explained an additional 3% variation (multiple correlation=0.65, \(R^2=0.42\)); and circulation explained less than 1% variation (multiple correlation=0.656, \(R^2=0.43\)). Consequently, for the sake of simplicity, one can calculate the estimated number of reference transactions by using operating expenditures alone. In such a case, the regression equation becomes:
Estimated number of reference trans. = 139.6 + 0.43 (total operating expenditures in $1,000s) ...(2)

The 95% confidence interval for the constant is 102.9 to 176.4 while that of the regression coefficient for the operating expenditures is 0.39 to 0.46. Using equation 2, one can compute the expected number of reference transactions for the New Jersey Institute of Technology since its library operating expenditures is known as $464,000. Computation would yield that the estimated number of reference transactions are 339.1 which is close to the number obtained by using equation 1. A regression plot between reference transactions and operating expenditures is shown in Figure 1. One can also use this plot to arrive at the estimated number of reference transactions for a given library operating expenditures.

ii. Directional transactions:

Similar procedure can be applied in estimating the expected number of directional transactions. Again, library operating expenditures, enrollment, and circulation are the best predictors for the directional transactions. The regression equation for the directional transactions is:

\[
\text{Number of directional transactions} = 132.0 + 0.18 \text{(Operating expenditures in}$ \$1,000s) + 0.036 \text{(Number of enrolled)} + 0.91 \text{(Number of items circulated in} \$1,000s)\]

...(3)
DATA WITH REGRESSION LINE AND C.I.

Fig. 1
Multiple correlation for this equation is 0.58 which means that nearly 34% of variation in directional transactions is explained by the combined effects of the three independent variables chosen. As in the case of reference transactions, all the remaining variables, together, explained much less than 1% variation in the directional transactions. Again, given the fact that library expenditures alone explained nearly 30% of the variation in the directional transactions, one can use the simple regression equation involving this variable alone to estimate the number of directional transactions. Such an equation is:

\[
\text{Number of directional transactions} = 198.4 + 0.48 \times \text{expenditures in $1,000s)}
\]

...(4)

Figure 2 is the graphical representation of this regression equation.

iii. Total number of transactions:

One can also use the regression analysis procedure to compute the expected number of all transactions at reference desks. All transactions at reference desks means the sum of reference and directional transactions reported by the libraires. The simple
equation for predicting this variable is:

\[
\text{Number of total transactions} = 333.9 + 1.08 \times \text{Library expenditures in $1,000s} \quad \ldots (5)
\]

This equation indicates that for every increase of $1,000 in the library operating expenditures, one can expect slightly over one transaction per typical week at reference desk. Figure 3 represents this regression equation.

iv. Regression forced through zero:

The regression equations 1 to 5 described so far suffer from one logical inconsistency. According to these equations, if all the independent variables are set to zero, we can still expect a certain number of reference and directional transactions. Equation 5, for instance, says that when the operating expenditures are zero, one should expect 334 transactions per typical week at reference desks. Since this is improbable, we need a regression equation which goes through zero when all the independent variables in the equation are set to zero. The simplest set of regression equations under this condition are:
\[
\begin{align*}
\text{Number of reference transactions} &= 0.49 \text{ (Library operating expenditures in } \$1,000\text{s)} \\
\text{Number of directional transactions} &= 0.57 \text{ (Library operating expenditures in } \$1,000\text{s)} \\
\text{Number of all transactions} &= 1.06 \text{ (Library operating expenditures in } \$1,000\text{s)}
\end{align*}
\]

V. Conclusions:

The regression equations presented in the analysis section are useful for quantitatively evaluating the reference service of individual libraries. Using one or more of these equations, any library can compute the expected number of reference and directional transactions and, then, compare these figures with the observed number of transactions. At this point, a note of caution should be sounded. First, the data on reference and directional transactions does not measure the quality of reference service. Second, the numbers obtained from the regression equations should be used only as guidelines and caution should be exercised in interpreting them. For instance, if
the observed number of all transactions is less than the expected number of transactions for a given library, it does not automatically follow that its reference service is poor compared to its peer institutions. The discrepancy may be explained, in addition to random statistical variations, by the fact that the reference department may be spending significant resources on library-use instruction or other comparable activities.

The regression equations developed so far could only explain less than 50% of the variation in the reference and directional transactions. This observation may mean two things. First, there could be other potentially useful independent variables but which are not included in the equations. One such variable that quickly comes to mind is the number of person-hours involved in the provision of reference service. LIBGIS data provides neither the number of reference librarians nor the number of hours the reference service is open to the public. The person-hour statistic could be more useful in explaining the variations in the reference and directional transactions than enrollment and circulation. Second, the linear regression model may be inappropriate and it is conceivable that the relationship between the dependent and independent variables could be nonlinear. However, preliminary
examination of some standard nonlinear relationships such as exponential relationship did not provide better explanation. But additional work could still be done in this area including the development of a theoretical model for reference service involving a number of independent variables.

Attempts were also made to derive regression equations for different types of libraries (e.g., public or private; university or four-year institution, etc.) by introducing dummy variables into the equations. This attempt also did not succeed because of two reasons. First, the dummy variables did not provide additional explanation for the variation in the reference and directional transactions. Second, the partial regression coefficients for dummy variables were not significant. This, again, confirms the earlier observation that the size of the library operating budgets, enrollment, and circulation are better able to explain the variations in reference and directional transactions than any other institutional information.
References:

1. Emerson, Katherine. Symposium on measurement of reference. RQ 14 (Fall 1974) 7-19

   (b) Ciucki, Marcella. Recording of reference/information service activities: A Study of forms currently used. RQ 16 (Summer 1977) 273-283


6. Definitions and Instructions attached to the College and University Library-Fall 1977 questionnaire which is a part of the Higher Education General Information Survey and Library General Information Survey conducted by the U.S. Department of Health, Education, and Welfare.