

Chapter 2: Heritage Health Index Methodology

prepared with RMC Research Corporation

Sampling Frame and Sample Selection

Heritage Preservation supplied RMC Research Corporation (RMC) with a sampling frame of over 35,000 collecting institutions (Heritage Health Index Development, p. 9). Heritage Preservation categorized each institution into one of five institution types: archives, libraries, historical societies, museums, or archaeological repositories/scientific research collections. In addition, Heritage Preservation designated each institution into one of three sample groups (Heritage Health Index Development, p. 6). The 1st Target Group included the largest and most significant collecting institutions of all types and in all U.S. states and territories; the 2nd Target Group included additional large or mid-sized institutions with important collections; and the Random Sample group included all remaining institutions to be sampled. The institutions in the 1st and 2nd Target Groups were sampled at 100%, whereas the type of institution was used as the primary stratum for drawing the sample from the Random Sample group. Institutions in this group were selected proportionately within type and within the state by zip code. The sampling strategy was designed to yield approximately 15,300 institutions.

Figure 2.1 shows the number of institutions in the sampling frame, the sampling rates applied, and the resulting number of sampled institutions within sample group and institution type.

Phone Verification

During the test of the Heritage Health Index, researchers found that direct phone contact with institutions to verify specific addresses, confirm appropriate contact names, and alert institutions about receiving the Heritage Health Index survey improved the likelihood of the survey getting to the right person in the institution and improved response rates. Heritage Preservation handled phone calls to all 1st Target Group institutions to verify contact information and brief the institutions on the intent and importance of the study.

RMC subcontracted with RKM Research and Communication in Portsmouth, New Hampshire, to verify contact information and eligibility for each institution in the 2nd Target and Random Sample Groups. Phone calls were made from July 11, 2004, to August 9, 2004. Phone verification resulted in the exclusion of 211 (1.5%) Random Sample institutions because of duplicate entries or the lack of collections. These institutions, along with those from the Random Sample that

Fig. 2.1 Sampling Rates and Number of Sampled Institutions (by sample group and type)

	Archives	Libraries	Historical Societies	Museums	Archaeological Repositories/ Scientific Research Collections	TOTAL SAMPLE n=
1st Target Group						
Sampling Rate	100%	100%	100%	100%	100%	
n=	70	179	39	163	35	486
2nd Target Group						
Sampling Rate	100%	100%	100%	100%	100%	
n=	42	449	16	306	60	873
Random Sample						
Population N=	429	14,010	4,288	12,154	1,897	
Sampling Rate	100%	35%	51%	35%	100%	
n=	429	4,908	2,204	4,273	1,897	13,711
TOTAL SAMPLE						
n=	541	5,536	2,259	4,742	1,992	15,070

asked not to participate in the study,¹ were removed from the survey mailing list. Seventeen of the 2nd Target Group institutions asked not to participate; however, Heritage Preservation wrote a letter asking them for their participation since they had been specifically selected and represented significant collections.

Figure 2.2 shows the numbers and percent of institutions found to be non-eligible for the study and the resulting number of institutions by sample group and type included in the survey mailing.

Web-based Survey Design

It was Heritage Preservation's intent to make participation in the Heritage Health Index study as easy and accessible as possible. A paper survey was designed along with an identical online Web-based survey. A consultant from Cultural Logic in Medford, Massachusetts, was contracted to design the Web survey and provide technical assistance during the data collection phase. Institutions were assigned unique passwords for access to their survey and to ensure data security. Cultural Logic and Heritage Preservation worked together to identify a comprehensive set of dynamic validation checks for online data entry. Consistency checks were developed across questions, and possible out-of-range values were defined for the numeric responses (e.g., operating budgets, number of units in a collection). Error

and warning messages were programmed into the online survey to ensure data integrity.

Information gathered from the Heritage Health Index test phase indicated that many institutions needed to gather information for the survey from several people or departments within their institution. With this in mind, the online survey was designed to allow for data entry over more than one online session, and indicators were built into the online system so users were aware of the section-by-section status of completion. Additional features included allowing respondents to print out the completed survey before final submission of the data and giving respondents access to some preliminary survey results.

Survey Mailing and Administration

All institutions were mailed a Heritage Health Index survey package on August 16, 2004. This package included an introduction letter signed by Heritage Preservation President Lawrence L. Reger and IMLS Director Robert S. Martin, Ph.D, the questionnaire booklet (Appendix F), instruction and Frequently Asked Question sheets, a unique online password, a flyer about accessing the Heritage Health Index online, and a return postage-paid envelope for those choosing to submit the survey on paper. All institutions were given contact phone numbers and e-mail addresses at Heritage Preservation and RMC as resources for content or technical questions.

Fig. 2.2 Results of Phone Verification and Number of Institutions in the Survey Mailing (by sample group and type)

	Archives	Libraries	Historical Societies	Museums	Archaeological Repositories/ Scientific Research Collections	TOTAL
Phone Verification						
Random Sample	429	4,908	2,204	4,273	1,897	13,711
Non-Eligible	- 9 (2%)	-28 (.6%)	-21 (1%)	-114 (3%)	-39 (2%)	-211 (1.5%)
Survey Mailing						
1st Target Group	70	179	39	163	35	486
2nd Target Group	42	449	16	306	60	873
Random Sample ¹	400	4,772	2,134	4,103	1,826	13,235
TOTAL	512	5,400	2,189	4,572	1,921	14,594

1. Surveys were not sent to Random Sample Group institutions that asked not to participate during phone verification.

Institutions were asked to submit data by October 12, 2004.

Contact information for mailed surveys that returned to RMC with “no such address” or “no longer forwarding” were sent to Heritage Preservation for further investigation. When possible, surveys were sent out a second time with the corrected contact or address.

A month after the surveys were mailed, reminder/thank-you postcards were sent to all institutions. Heritage Preservation made personal reminder calls to all 1st and some 2nd Target Group institutions. The survey was administered using all the steps tested by Aeffect, Inc., and Heritage Preservation (Heritage Health Index Development, p. 5). However, because of budget limitations and to encourage as many online survey responses as possible, a second copy of the paper survey was not distributed. Instead, a personal reminder letter that included each institution’s unique password was sent to all institutions that had not yet responded. Several weeks later, RMC sent out two final e-mail reminders to those who had not responded at all or who had online surveys in progress and for whom Heritage Preservation had e-mail addresses. Heritage Preservation allowed two extension dates, with a final cut-off of December 15, 2004.

Data was collected from mid-August 2004 until December 15, 2004, from online entries and from RMC staff who entered paper survey returns via the Web site. All data passed through the quality control data checks within the online data entry system. Almost three quarters (73%) of the institutions chose to submit data using the Web-based survey. Libraries were most likely (81%) to respond using the online survey and historical societies were the least likely (54%).

Final Dispositions and Response Rates

Since the original Heritage Health Index sampling frame was derived from mailing lists, cultural institutions and departments within an institution could appear on the list. The original sample unit was defined as a cultural institution’s address or location. To encourage participation and obtain the most accurate data possible, the study allowed institutions to define their collection entity, which resulted in redefining the

sample unit. In many cases, one institution was responsible for the preservation of collections both at their location and elsewhere. Sometimes it was easier for one institution to respond to the survey for itself and other affiliations. Institutions covered under another institution usually identified themselves to RMC by e-mail, written note, phone call, or online survey comments. The sampled institutions already represented in another sample unit were tracked and ultimately given the disposition of “non-sample.” That is, they represent a percentage of institutions that should be removed from the study population.

Other “non-sample” institutions were also tracked; this included duplicate entities, institutions with no holdings for which they take a preservation responsibility, and institutions that had ceased operation. Heritage Preservation was able to identify every sampled institution in the 1st and 2nd Target Groups, whether they responded or not, by their sample eligibility. On the other hand, at the end of the data collection, it was unknown whether 75% of the Random Sample Group was eligible to be included in the sample. Using the eligibility findings for each of the five institutional types for the Random Sample, the unknown sample group was reduced by the non-eligible percent.

Figure 2.3 shows a) number of institutions mailed to; b) number of institutions with unknown eligibility after data collection; c) percent of institutions found to be non-eligible (applied to b) for sample reduction); and d) estimated number of eligible institutions in the sample. The revised eligible sample includes all respondents, any non-respondents known to be eligible, and the reduced number of non-respondents of unknown eligibility.

Over 3,600 institutions responded to the Heritage Health Index survey. However, after examining the data, 7% of those survey responses were not considered complete. The majority of those cases were from respondents who entered their data online but never returned to the survey to complete Section F on the quantity and condition of collections. Heritage Preservation and RMC decided to drop those cases from the response rates and from the analyses.

Of the 13,590 eligible institutions, 3,239 com-

Fig. 2.3 Final Dispositions from Survey Mailing (by sample group and type)

	Archives	Libraries	Historical Societies	Museums	Archaeological Repositories/ Scientific Research Collections	TOTAL
1st Target Group						
a. Survey Mailing	70	179	39	163	35	486
b. Unknown Disposition	0	0	0	0	0	0
c. % Found Non-Eligible	(8.6%)	(6.1%)	0	0	(5.7%)	(3.9%)
d. Eligible Sample	64	168	39	163	33	467
2nd Target						
a. Survey Mailing	42	449	16	306	60	873
b. Unknown Disposition	0	0	0	0	0	0
c. % Found Non-Eligible	(2.4%)	(0.7%)	0	(1.3%)	0	(0.9%)
d. Eligible Sample	41	446	16	302	60	865
Random Sample						
a. Survey Mailing	400	4,772	2,134	4,103	1,826	13,235
b. Unknown Disposition	275	3,756	1,555	2,999	1,316	9,901
c. % Found Non-Eligible	(3%)	(4%)	(6%)	(11%)	(12%)	(2%)
d. Estimated Eligible Sample	387	4,594	2,016	3,657	1,604	12,258
TOTAL SAMPLE						
a. Survey Mailing	512	5,400	2,189	4,572	1,921	14,594
d. Estimated Eligible Sample	492	5,208	2,071	4,122	1,697	13,590

Fig. 2.4 Response Rates (by sample group and type of institution)

	Archives	Libraries	Historical Societies	Museums	Archaeological Repositories/ Scientific Research Collections	TOTAL
1st Target	92%	90%	90%	90%	91%	90%
2nd Target	41%	39%	31%	54%	45%	45%
Random Sample	25%	18%	21%	20%	22%	20%
TOTAL	35%	22%	22%	25%	24%	24%

Fig. 2.5 Response Rates (by sample group and region)

	Northeast	Mid-Atlantic	Southeast	Midwest	Mountain-Plains	West	TOTAL
1st Target	92%	89%	88%	92%	94%	89%	90%
2nd Target	59%	39%	42%	50%	49%	35%	45%
Random Sample	20%	18%	19%	20%	20%	20%	20%
TOTAL SAMPLE	25%	23%	23%	23%	24%	24%	24%

pleted the Heritage Health Index survey, providing an overall response rate of 24%. The institutions within the 1st Target Group, which represent the largest collections in the country, had a 90% response rate. Figures 2.4 and 2.5 show that generally, the 1st Target Group and the Random Sample Group responded similarly across institutional types and geographic regions. However, the 2nd Target Group had over a 20% difference in response rates across types and regions.

Test Surveys and Volunteer Respondents

Heritage Preservation made all test surveys available to RMC. These surveys were added to the Heritage Health Index data file, along with 20 surveys from institutions that were not in the sample but asked to participate.

Figure 2.6 below shows the total number of par-

ticipating institutions by sample group and type. The distributions across type are presented for the participants and the revised eligible sample. Note the participants closely represent the sample by within two percentage points.

Data Cleaning

The survey data was originally stored in an ACCESS database. It was imported into Statistical Package for the Social Sciences (SPSS); both data sources served as platforms for the data cleaning process.

RMC and Heritage Preservation reviewed any questionable data. Numeric outliers were checked and assigned as missing if left unresolved. Any question that allowed for “other” as a response was blindly reviewed by Heritage Preservation. Examining “other” responses to Section F on the

Fig. 2.6 Number of Participant and Sample Institutions (by sample group and type) and Percent of Participant and Sample Institutions (across type)

	Archives	Libraries	Historical Societies	Museums	Archaeological Repositories/ Scientific Research Collections	TOTAL
1st Target	59	150	35	145	30	419
2nd Target	17	174	5	164	27	387
Random Sample	96	832	418	731	356	2,433
Test	7	37	13	43	11	111
Volunteers	1	3	5	5	6	20
TOTAL	180	1,196	476	1,088	430	3,370
Percent across Type	5%	36%	14%	32%	13%	100%
Estimated Eligible Sample	492	5,208	2,071	4,122	1,697	13,590
Percent across Type	4%	38%	15%	30%	13%	100%

quantity and condition of collections was a considerable task. Heritage Preservation made decisions to reassign “other” responses to the appropriate collections subcategories, created a new subcategory to capture the response, or approved the “other” response. Data had to be added, deleted, or readjusted for more than 500 responses. Heritage Preservation reviewed all comments made in the option question (G10) that asked institutions to explain their most pressing conservation/preservation needs and categorized the responses. RMC applied all data revisions to the SPSS data files.

Data Analysis

Heritage Preservation staff and RMC Research Corporation reviewed the Heritage Health Index data and made initial decisions regarding data tabulation. In March 2005, Heritage Preservation presented the findings to senior IMLS staff and discussed data analysis. In May 2005, Heritage Preservation convened a group of 14 conservators and other collections professionals for their feedback on analysis and input on identifying the survey’s key findings. This group included members of the Heritage Health Index Working Groups that helped develop the survey instrument, individuals that completed the survey for their institution, and several individuals who did not have prior contact with the project and could provide a fresh perspective. From the close of data collection until the publication of this report (December 2004 to December 2005), Heritage Preservation staff consulted with its board members and Working Group members to discuss questions raised by the data analysis.

Assignment of Institution Type and Self-Reported Institution Type

For sampling purposes, Heritage Preservation categorized all institutions into one of five types: archives, libraries, historical societies, museums, and archaeological repositories/scientific research collections. For the most part, these assignments were made based on the institutions’ names. The survey question B1 asked institutions to choose from 21 possible categories to best describe the primary function or service of their institution. All analyses performed used the self-identified institution type from the questionnaire.

Figure 2.7 compares the original type assignment and the self-reported type. Museums were the most likely to be misclassified while establishing the sampling frame, possibly because museum was in the organization’s name but was not its primary function.

Subgroup—Institution Type

Heritage Preservation initially viewed the Heritage Health Index data by institutional type as defined by the list of 21 types of institutions in question B1, which asked participants to select their primary function or service. However, viewing the data by these many categories was cumbersome and, in the case of some groups (e.g., children’s museums, arboretums, aquariums), insignificant because the data was based on few responses (Characteristics of Collecting Institutions in the United States, figure 3.1, p. 23). Heritage Preservation identified types of institutions that had similar findings and whose data could be aggregated. In consultation with IMLS staff, Heritage Preservation narrowed the list of

Fig. 2.7 Percent of Respondents for Assigned and Self-Identified (by type)

	Archive	Libraries	Historical Societies	Museums	Archaeological Repositories/ Scientific Research Collections	TOTAL
Assigned Type n=3,370	5%	36%	14%	32%	13%	100%
Self-Identified Type n=3,370	6%	35%	11%	41%	7%	100%

21 institutional types to 10:

1. archives
2. public libraries
3. special libraries (e.g., law, hospital, and religious libraries and libraries for the blind and handicapped)
4. academic libraries
5. independent research libraries (includes national and state libraries)
6. historical societies
7. art museums
8. history museums/historic sites/other museums (includes historic houses/sites, history museums, living history museums, general museums, specialized museums, children's museums)
9. science museums/zoos/botanical gardens (includes natural history museums, science/technology museums, nature centers, planetariums, observatories, arboretums, botanical gardens, aquariums, zoos)
10. archaeological repositories/scientific research collections (institutions that would not be classified as museums by IMLS's definition).

Several survey questions included the answer choice of "other, please specify." These "other" responses were incorporated into existing answer categories. Doing so was particularly important in the case of question B1 on primary function since the survey data was to be reviewed by institutional type.

For a broader view by institutional type, Heritage Preservation grouped the institutions into the five groups by which the survey sample was stratified: archives, libraries, historical societies, museums, and archaeological repositories/scientific research collections. Viewing the data by five types of institution rather than ten types results in a lower margin of error, and so it is used most frequently in the Heritage Health Index reporting. In some instances, it is useful to view data by specific institutional type.

Subsidiary Functions: Archives and Libraries

Heritage Preservation grappled with how to capture data on archives, as they are often subsidiaries of other institutions. Survey participants were instructed to complete the survey for all collections at the institutions, and the exam-

ple of a subsidiary archives or library was used. The questionnaire asked institutions to identify one primary function or service and to select as many secondary functions or services as applicable. Therefore, the Heritage Health Index data may be viewed by institutions that selected archives as a primary function (referred to in this report as "stand-alone archives") and by institutions that indicated archives as a primary or additional function.

Since the number of stand-alone archives in the Heritage Health Index survey population was small, this group was sampled at 100% in the stratified sample. In total, 180 stand-alone archives participated in the survey and their data has a margin of error of 5.5%, a slightly higher rate than data viewed by other types of institutions. Institutions that indicated archives as a secondary function totaled 44%, the most frequent secondary function. The second most common secondary function was "library," with 22% indicating they had this additional purpose. These subgroups are referenced in the report when they further illuminate Heritage Health Index findings.

Subgroup—Institution Size

Budget and collection size data were reviewed to categorize institutions by actual size. When available, Heritage Preservation adapted definitions of size from other professional associations' publications or surveys to make the Heritage Health Index as comparable to other studies as possible. The definitions were reviewed and approved by IMLS staff and other project advisors. As explained in the chart below, size definitions use different criteria based on type of institution.

Archives

The size of archives was based on the quantity of unbound sheets and other archival materials reported in the Heritage Health Index. For some archives, significant photographic, moving image, or recorded sound collections were taken into account as well as unbound sheets.

Large	-more than 5,000 linear feet of unbound sheets
	-institutional budgets that are appropriate for a larger institution

	-all National Archives and Records Administration facilities and most state archives
Medium	-1,000-4,999 linear feet of unbound sheets -reasonable institutional budget size (unusually small budgets may have resulted in reclassification as small) -remaining state archives
Small	-fewer than 1,000 linear feet of unbound sheets

Academic and Independent Research Libraries

The size of academic and independent research libraries was based on the total volume holdings of respondents as reported in the 2004-2005 *American Library Directory*. Libraries reporting significant holdings in unbound sheets may have been moved to a larger category.

Large	-more than 1,500,000 total volume holdings -all members of the Association of Research Libraries -most state libraries
Medium	-250,000-1,499,999 total volume holdings -all members of the Oberlin Group of Liberal Arts College Libraries -remaining state libraries
Small	- fewer than 250,000 total volume holdings

Public Libraries

The size of public libraries was based on the service populations of respondents as reported in the 2004-2005 *American Library Directory*. Definitions are based on those used by the Public Library Association.

Large	- service population 100,000 or greater
Medium	- service population 25,000-99,999
Small	- service population less than 25,000

Archaeological Repositories

The size of archaeological repositories was based on the quantity of individually and/or bulk cataloged archaeological collections as reported in the Heritage Health Index. If the collection size was not provided, decisions were made on institutional

budget size, with large=more than \$1,000,000, medium=\$350,000-\$999,999, small=less than \$350,000, except in the case of labs that offered archaeological services, which often have larger budgets. In this case, additional research was done to determine number of holdings.

Large	- more than 500,000 individually cataloged archaeological collections and/or more than 5,000 cubic feet of bulk archaeological collections
Medium	- 5,000-499,999 individually cataloged archaeological collections and/or 1,000-4,999 cubic feet of bulk archaeological collections
Small	- fewer than 5,000 individually cataloged archaeological collections and/or fewer than 1,000 cubic feet of bulk archaeological collections

Agency or university collection with scientific specimen artifact collections with herbarium and/or zoological focus

The size of these collections was based on the quantity of botanical specimens and/or zoological specimens as reported in the Heritage Health Index.

Large	- more than 500,000 botanical specimens and/or zoological specimens
Medium	- 50,001-500,000 botanical specimens and/or zoological specimens
Small	- 50,000 or fewer botanical specimens and/or zoological specimens

Agency or university collection with scientific specimen artifact collections with geological or paleontological focus

The size of these collections was based on the quantity of geological or paleontological specimens as reported in the Heritage Health Index.

Large	- more than 200,000 geological and/or paleontological specimens
Medium	- 10,001-199,999 geological and/or paleontological specimens
Small	- 10,000 or fewer geological and/or paleontological specimens

Museums

The size of museums was based on the institutional budget as reported in the Heritage Health

Index. Definitions are based on what the American Association of Museums had used in several reports, such as the 1989 *National Museum Survey* and the biennial AAM Museum Financial Information surveys (last used in the 1999 study). Dollar figures, not updated since 1989, have been adjusted for inflation. If institutional budget information was not provided for museums, the 2005 *Official Museum Directory* was consulted for staff size and used to place museums in size categories with large=more than 10 full time paid staff, medium=4-10 full time paid staff, and small=3 or fewer full time paid staff.

Aquariums, Zoos

Large - institutional budget more than \$4,500,000
 Medium - institutional budget \$1,500,000-\$4,500,000
 Small - institutional budget less than \$1,500,000

Arboretums, Botanical Gardens, Art Museums, Children’s Museums

Large - institutional budget more than \$1,500,000
 Medium - institutional budget \$300,000-\$1,500,000
 Small - institutional budget less than \$300,000

General Museums, Historic House/Sites, History Museums, Historical Societies, Specialized Museums

Large - institutional budget more than \$1,500,000

Medium - institutional budget \$500,000-\$1,500,000
 Small - institutional budget less than \$500,000

Natural History Museums

Large - institutional budget more than \$1,500,000
 Medium - institutional budget \$350,000-\$1,500,000
 Small - institutional budget less than \$350,000

Nature Centers, Planetaria

Large - institutional budget more than \$1,200,000
 Medium - institutional budget \$350,000-\$1,200,000
 Small - institutional budget less than \$350,000

Science/Technology Museums

Large - institutional budget more than \$7,500,000
 Medium - institutional budget \$1,500,000-\$7,500,000
 Small - institutional budget less than \$1,500,000

Subgroup—Region

The survey sample was stratified by state to ensure accurate geographical representation. However, even with a large survey sample (almost half the entire sampling frame), the relatively few number of institutions in some states would have

Fig. 2.8 Definitions of Region

Region	States and Territories Included
Northeast	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
Mid-Atlantic	District of Columbia, Delaware, Maryland, New Jersey, New York, Pennsylvania, Puerto Rico, U.S. Virgin Islands
Southeast	Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia
Midwest	Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Ohio, Wisconsin
Mountain-Plains	Colorado, Kansas, Montana, Nebraska, New Mexico, North Dakota, Oklahoma, South Dakota, Texas, Wyoming
West	Alaska, American Samoa, Arizona, California, Guam, Hawaii, Idaho, Nevada, Northern Mariana Islands, Oregon, Utah, Washington

required sampling at 100% and response rate of close to 100% to produce reliable data by state. Therefore, Heritage Preservation decided that obtaining results by state would not be feasible. Several states with large populations of collecting institutions did achieve response rates that allowed their data to be reported with relatively low margins of error, and this data could be further analyzed in the future. Additional data collection by state could also be done to produce Heritage Health Index results by state. In this report, geographic results are reported by region using definitions from the American Association of Museums' *Data Report from the 1989 National Museum Survey* (see figure 2.8).

Subgroup—Governance

The Heritage Health Index survey question B5 asked institutions to indicate the governance under which they operate. Responses included “college, university, or other academic entity,” “nonprofit, non-governmental organization or foundation,” “corporate/for-profit,” “federal,” “state,” “local (county or municipal),” and “tribal.” Institutions operated by an academic institution were asked to complete a follow-up question (B6) on the governance of the institution; answer choices were “private college or university,” “state college or university,” “county or municipal college or university,” or “other, please specify.” When analyzing the data by governance, private college/university results were combined with private nonprofit, state college/university were combined with state governance, and county/municipal college/university were combined with county or municipal governance. Data was also run by

academic institutions alone, and these results are mentioned in the report when relevant.

The governance question identified that some corporate/for-profit institutions replied to the survey, although this type of institution was not included in identifying the Heritage Health Index survey population. While data from this group was not dropped, corporate/for-profit is not identified when results are viewed by governance since they represent only 1% of surveyed institutions.

Study Projections

The survey data has been weighted to produce estimates that reflect the defined population of U.S. collecting institutions, including types of institutions, sizes of institutions, and regional locations. There was no evidence to show that the non-respondents from the 2nd Target Group or the Random Sample Group have any characteristics different from the respondents. Since the responding institutions from the 1st Target Group differ widely in the types and sizes of collections they hold and because this group received a high response rate of 90%, it was decided not to include these non-respondents in the population projections. In addition, any collecting institution that did not appear in the original sampling frame would not be accounted for in the projected population. With this in mind, the population estimate is conservative and excludes 10% of the large institutions, which hold significant collections.

To produce correct population estimates for the Heritage Health Index, weights were applied to the sample data to compensate for the following:

Fig. 2.9 Percent and Number of Institutions for Survey Respondents and Estimated Population (by type)

	Archives	Libraries	Historical Societies	Museums	Archaeological Repositories/ Scientific Research Collections	TOTAL
Sampled	6%	35%	11%	41%	7%	100%
Survey Respondents	n=206	n=1192	n=361	n=1384	n=227	n= 3,370
Estimated Population	3%	43%	11%	39%	4%	100%
After Weighting	N=1,033	N=13,324	N=3,303	N=12,057	N=1,110	N=30,827

- Differential probabilities of sample selection for institutional types and sample groups; and
- Non-respondents in the 2nd Target Group and the Random Sample Group.

The base weight for each institution is equal to the reciprocal of its probability of selection for the sample group and institution type. The base weights were adjusted for non-respondents by subgroup to ultimately produce the study population of collecting institutions.

The sample surveys of 3,370 institutions represent nearly 31,000 collecting institutions. The proportion of sampled institutions by type has been adjusted by the weighting to accurately reflect the population proportions. The largest adjustment was made to the subgroup of sampled libraries, where they were underrepresented by eight percent. The results of the weighting scheme are presented in figure 2.9.

Overall Confidence Intervals

To describe the precision of institutional estimates made from the Heritage Health Index survey, a level of confidence (or margin of error) was calculated.² The 95 percent confidence level for all institutions is generally no greater than +/- 1.5 percentage points around any given percent reported. Examining results by institution size produces margins of error less than 3.5 percentage points. Examining results by the five institutional types has confidence intervals ranging from 2.3 to 5.5 percentage points. The confidence intervals are larger around estimates for the smaller subgroups, such as by the ten institu-

tional types or when data is reported by size *and* type of institution. The margins of error for the overall sample and by type, by size, and within type and size are presented in figure 2.10.

Methods for Weighting or Imputing Data

As with most surveys, both unit (institution) and item (question) non-response is unavoidable. Weighting adjustments were made for unit non-response as described previously. This section focuses on item non-response and the method used to compensate for missing responses to questions.

Annual Operating Budgets

To estimate total annual operating budgets, total budgets for conservation/preservation, and the percentages spent on preservation for the population, additional weights needed to be applied. The reporting of financial data had a higher level of non-response than other questions, and the additional weights compensate for those institutions not reporting. The variability in budgets is very high across all institutions but much smaller for subgroups when defined by size and type. Institutions were cross categorized by sizes (large, medium, small) and 10 types (archives, public libraries, academic libraries, independent research libraries, special libraries, historical societies, art museums, history museums/historic sites/other museums, science museums/zoos/botanical gardens, archaeological repositories/scientific research collections). Each institution with valid data for operating budget and conservation

Fig. 2. 10 Margins of Error (by type, by size, and within type and size)

	Archives	Libraries	Historical Societies	Museums	Archaeological Repositories/Scientific Research Collections	TOTAL
Large	6.1	8.8	4.8	3.8	10.4	2.6
Medium	13.3	21.0	5.1	5.5	10.4	3.4
Small	9.2	5.0	3.6	3.2	7.1	2.0
TOTAL	5.5	4.6	2.6	2.3	5.2	1.5

2. The margin of error was calculated using the following formula: $1.96 \cdot \sqrt{[(.5 \cdot .5) / n]} \cdot [(N - n) / N - 1]$ where n assumed 100% item response rate.

budget was weighted according to its cross classification and proportion of missing data.

Quantity and Condition of Collections

Institutions were asked to report on the number of collection items they hold and the condition of the items for more than 50 different types of collections. Many institutions reported holding specific types of collections but were unable to report on the quantity or did not respond to the conditions. Since one of the main objectives of the Heritage Health Index was to report on the conditions of all collection items in the United States, missing data was imputed with values from similar institutions.

The methods used for imputing quantity and condition data were generally the same. However, it was the median quantity and the mean condition that was imputed. The variance in quantity size was so large that imputing the mean quantity would result in skewing the population estimates. Each of the 58 types of collections was isolated, and only those institutions holding that collection were aggregated by institution size (large, medium, small) and 10 types (archives, public libraries, academic libraries, independent research libraries, special libraries, historical societies, art museums, history museums/historic sites/other museums, science museums/zoos/botanical gardens, archaeological repositories/scientific research collections). The median

quantity of collection items and mean conditions were calculated for each possible subgroup. If a subgroup contained less than eight institutions contributing to the median or means, that subgroup was combined with another subgroup of the same size and similar type. For example, if there were fewer than eight small science museums reporting quantities of “art on paper,” their responses would be combined with another subgroup like small history museums/historic sites/other museums or all small museums.

If an institution indicated holding a specific type of collection but reported the quantity unknown, the median value by subgroup was imputed. If an institution indicated holding a specific type of collection but did not report on the conditions of the collection, the mean conditions by subgroup was imputed. The percentage of a collection in unknown condition was considered a valid response.

Rounding

RMC reported data to the first decimal place. In the Heritage Health Index report, all data is rounded to the nearest whole number. For results less than 0.5%, the number is displayed as zero. Due to rounding, responses may add to 99% or 101% rather than 100%. Data will not add to 100% in questions where multiple responses were allowed, and this is noted when it occurs.