What software do you use? Results from a survey of the APHEO membership

James Lane, MSc
Health Data Analyst, Durham Region Health Department

Release date: November 2016

Introduction

Within the Association of Public Health Epidemiologists (APHEO), day-to-day work relies heavily on a collection of computer software such as statistical software, survey tools, geographic information system (GIS) packages and data visualization programs. These tools come in a variety of versions and are published by multiple vendors. For all the differences within a given class of software, these tools are leveraged to perform many of the same tasks, leaving one to question the possible opportunity for collaboration with the goal of increased efficiency. Unfortunately, there is a lack of published data on the software in use by APHEO members, making the process of discovering like-minded users difficult. To begin to address this data gap, a survey of the APHEO membership was conducted with the goal of describing the software that APHEO members use.

Method

All members of APHEO were invited to participate in an online survey, using the APHEOlist listserv to distribute the advertisement and survey link. The survey remained open until the majority of public health units (PHU) responded including representation of the small, rural and northern public health units. In total, 81 responses from 40 distinct APHEO member organizations (32 public health units and eight non-public health unit organizations) were obtained.

The survey was created using Ocean by CognisantMD and contained a series of close-ended questions which asked responded to select the software they use at work. In addition, open-ended responses allowed participants to specify additional software beyond the pre-defined options.

Statistical analysis was completed using SAS version 9.4. Descriptive statistics including counts and proportions were calculated for all survey items.

Results

The survey received 81 responses with 69 respondents indicating they work at a public health unit and 12 indicating they do not work at a public health unit (see table 1 below). The survey received a response from each PHU with the exception of Elgin-St. Thomas, North Bay Parry Sound District, Peterborough County-City and Thunder Bay District. Non-public health unit organizations included Alberta Health Services, BORN Ontario, Cancer Care Ontario, HCM Group, Public Health Ontario and the Royal Ottawa Hospital Group.

Table 1. Respondents to the APHEO software survey by organization type (n=81)

Organization	Frequency	Per cent
Public Health Unit	69	85
Non-Public Health Unit	12	15

As seen in figure 2 below, in terms of the availability of statistical software, the majority of respondents indicated that they have access to Stata (69 per cent), followed by SPSS (60 per cent), SAS (32 per cent) and R (25 per cent). Likewise, 38 per cent of respondents indicated they have access to a single statistical software package; 40 per cent have access to two packages; and 22 per cent have access to three or more statistical packages.

Statistical software

Figure 1. Statistical software available to respondents at work (n=81; groups are not mutually exclusive)

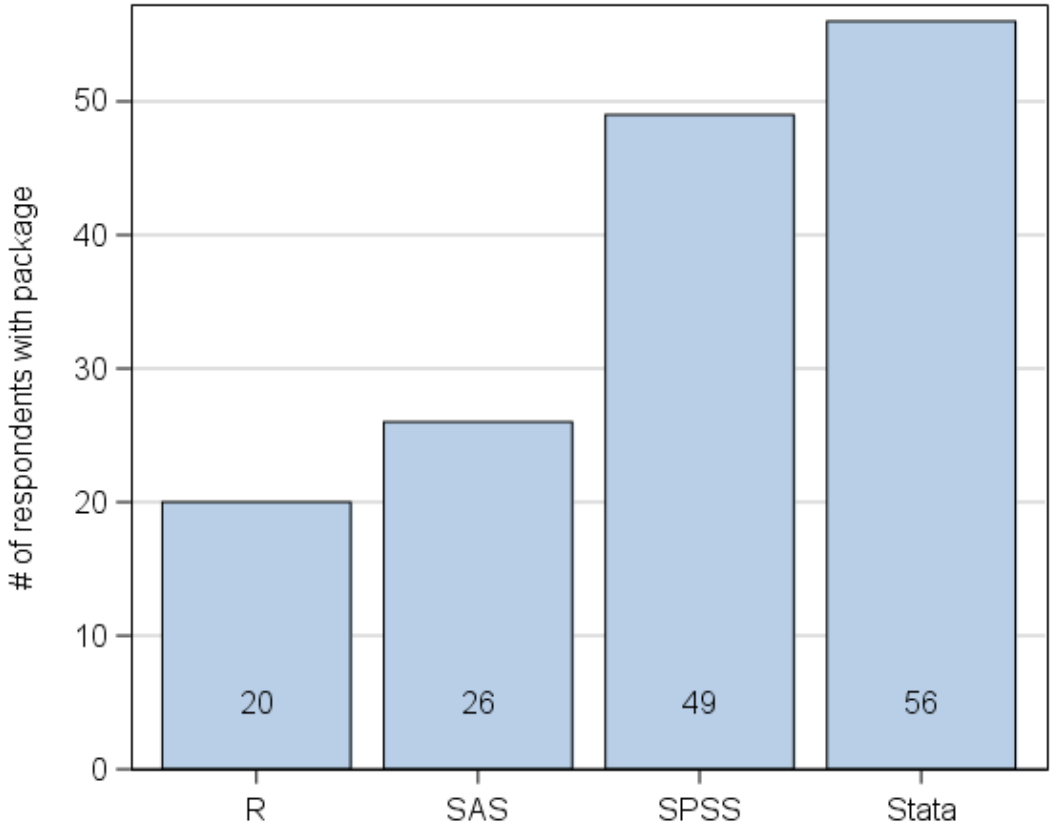

Table 2. Statistical software available to respondents at work (n=81; groups are not mutually exclusive)

Statistical Software	Count	Per cent
R	20	25
SAS	26	32
SPSS	49	60
Stata	56	69

Survey participants were asked to select the statistical software they use as their primary data analysis tool and the results are summarized in figure 2. Most respondents selected Stata as their primary data analysis tool (43 per cent); SPSS was selected by 26 per cent of respondents; and SAS by 20 per cent of respondents (see figure 2 below).

Figure 2. Statistical software package selected by respondents as their primary data analysis tool (n=81)

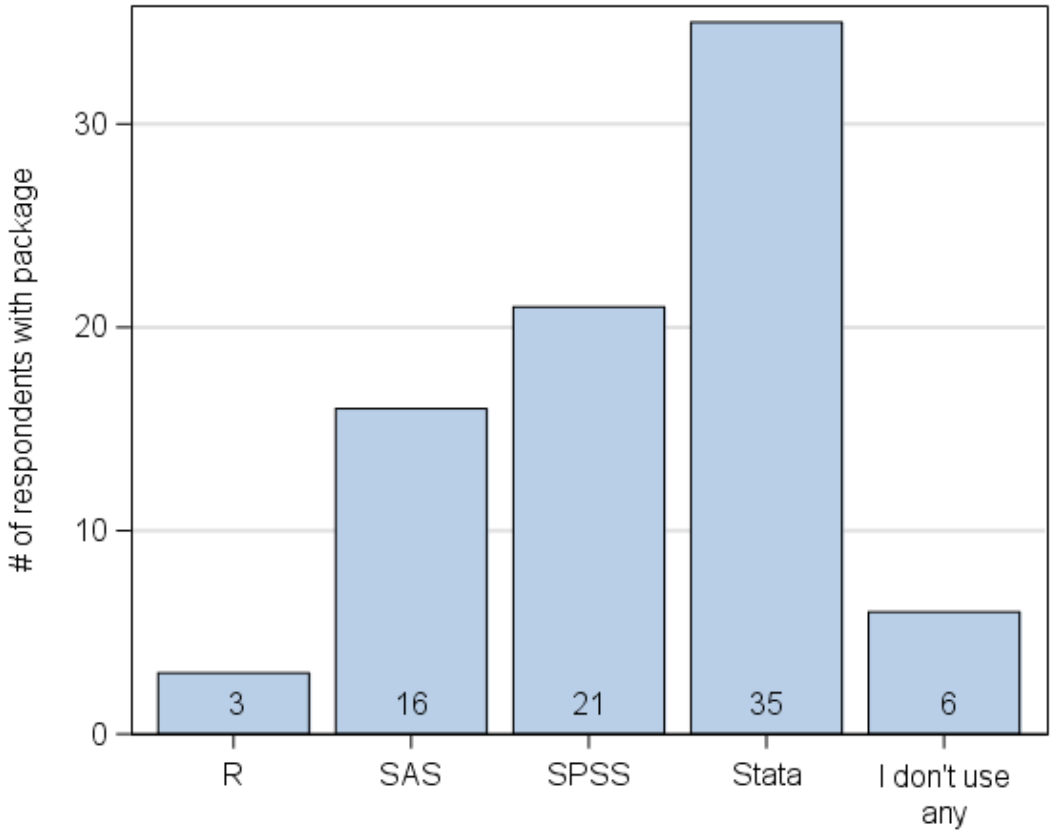

Table 3 Statistical software package selected by respondents as their primary data analysis tool (n=81)

Statistical Software	Count	Per cent
R	3	4
SAS	16	20
SPSS	21	26
Stata	35	43
I don't use any	6	7

Survey tools

As seen in table 4 below, the use of survey tools did not vary greatly across survey respondents. Approximately two-thirds of respondents indicated they use Fluid Survey as their survey tool (63 per cent), 9 per cent indicated they use either Cognisant MD or Simple Survey, while 21 per cent noted they use a survey tool other than the pre-

defined options. Responses in the group included Lime Survey, Remark Web Survey, Select Survey, Survey Monkey and iSurvey (data not shown).

Table 4 Survey tool selected by respondents as their solution for creating surveys (n=81)

Survey Tool	Count	Per cent
Cognisant MD	3	4
Fluid Surveys	51	63
I don't know	6	8
Simple Survey	4	5
Something else	17	21

Data visualization software

The results of the item asking survey respondents about data visualization tools are shown in table 5 below. Most survey respondents indicated they do not use a software tool for data visualization (63 per cent). Tableau was reported most frequently among those that indicated that they used a data visualization tool (16 per cent), followed by Microsoft Excel (7 per cent).

Table 5. Software selected by respondents as their data visualization tool (n=81)

Data Visualization Software	Count	Per cent
Microsoft Excel	6	7
Microsoft Power BI	4	5
Tableau	13	16
Something else	7	16
I don't use any	51	63

GIS software

As for the use of Geographic Information System (GIS) software, the majority of respondents indicated that they use ArcGIS in their work (60 per cent) while a third (33 per cent) of respondents indicated they do not use any GIS software (see table 6 below).

Table 6. Software selected by respondents as their GIS tool (n=81)

GIS Software	Count	Per cent
ArcGIS	49	60
QGIS	2	2
Something else	3	4
I don't use any	27	33

Discussion

The software tools used by APHEO members show similarities and differences. While the use of survey and GIS tools showed consistency, the use of statistical analysis software varied between respondents. This is noteworthy as statistical analysis represents a large proportion of the work carried out by APHEO members, and therefore offers the greatest opportunity for collaboration. The financial and person-time

cost of acquiring statistical software is a possible explanation, however the consistency reported for survey and GIS software provides evidence against this hypothesis.

With respect to data visualization tools, the results revealed that this activity is likely not a part of the duties of most APHEO members. This is not unexpected, given the relative infancy of data visualization as a medium and the barrier of learning a new skill that this task presents. However, as individuals change the way they consume information, and the scope and tools related to data visualization evolve, finding innovative ways to present information may become an important concept for public health epidemiologists.

Conclusion

This survey was, to the knowledge of the author, the first to collect information on the software tools used by APHEO members. The strengths of this survey included a near census of Ontario public health units with responses from 33 of 37 organizations. The limitations include the lack of a sampling strategy which may have introduced a bias towards more motivated individuals; and the use of the APHEOlist listserv to distribute the survey invitation, as certain APHEO members are more likely to keep up to date with this channel.

APHEO members use similar tools to accomplish similar tasks, presenting the opportunity to share knowledge between like-minded users. This sharing of knowledge has the ability to increase the comparability of results between users and decrease the likelihood of calculation errors. Further, sharing the expertise present within the APHEO membership can increase the efficiency with which tasks are completed and raise the collective knowledge across members. All of which carry the vision of APHEO, by advancing the professional practice of epidemiology in Ontario public health units.