

MaRS 2015 Venture Client Annual Survey - Methodology

OCTOBER 2016

DISTRIBUTION OF VARIABLES

Several different types of data are collected in the annual survey. These include:

- Likert scale variables. Typically these are opinion questions set on a 1-5 scale. Examples include "Likelihood to recommend MaRS services," "Importance of MaRS services," and "MaRS Impact on success."
- Categorical variables. Clients are asked to indicate which group they belong to among a set of mutually exclusive options. Examples include "What is the status of your venture? (Active / Ceased / Acquired)" or "What are the most important future services that MaRS could provide?"
- Binary variables, such as yes/no questions. For example, determining if their business was founded to address social or environmental impact as part of its core mission and vision?
- Quantitative questions. Clients were asked to fill in a number corresponding to their employment, revenue, funding or year incorporated.
- Open-ended questions. A "fill in the blank" box asking for a free-form response. Typically used to solicit additional opinions and feedback not captured in closed-end questions (e.g., "Please explain your answer on the impact that MaRS Discovery District has had on your venture's chance of success?")

SOFTWARE AND LOGISTICS

- The survey was administered by MaRS Data Catalyst, a distinct and separate business unit within MaRS that ensured that survey methodology best practices were adhered to and that surveys were sent directly to and responded directly from ventures. In doing so, it ensured that the survey adheres to provincial auditing guidelines.
- The survey was sent by email to 1,120 ventures on March 29, 2016; the deadline to respond was April 21, 2016. Respondents were asked to report data for the one-year period ending December 31, 2015.
- The survey was administered with FluidSurveys, a subsidiary of the larger SurveyMonkey, a commercial survey research tool employed by many large organizations.
- Preliminary client data pulled from SalesForce was used to pre-populate certain fields in the survey. The fields chosen were those where MaRS already had information that does not change from year to year, or if it does, not very frequently. The respondent could then confirm the information without having to enter

it again year-over-year. The pre-populated fields included: main contact information for the company, demographic data, incorporation date and postal code.

- After the survey was completed, the data was extracted, loaded and stored in a Microsoft SQL Server database hosted on a server maintained by Data Catalyst and their IT department. Although MaRS and Data Catalyst share a common IT subcontractor, non-Data Catalyst staff do not have direct access to the Data Catalyst servers. The analysis was conducted with Tableau software, a visual business intelligence and charting tool as well as “R”, a statistical programming language.

DEMOGRAPHICS

Demographics are reported at the level of the venture, rather than the individual. Because ventures can have multiple founders, this was asked as a single variable indicating the overall composition of the company. For reporting, rather than reporting the number of men and number of women founders, we reported the number of companies with all male founders, all female founders, or a combination of male and female founders. A similar method was applied to the other demographic questions: nationality, youth (<30 years old) and startup experience.

EXTRAPOLATION

Because the survey is not a random sample of our population, we are not able to assume that non-respondents are similar to respondents. Additionally, the distribution of the key variables of interest for extrapolation - revenue, funding and jobs - follows a "power law" or "long tail" shape, where a very small number of ventures are responsible for a majority of the substantial values. Because outliers drive the aggregate totals, we cannot extrapolate these out to a larger population.

RESPONSE RATES

The response rate of 51% is excellent for web surveys, far exceeding the typical 20-30% range achieved by general-interest surveys and polling. FluidSurveys reports that 25% is the typical response rates for email surveys with >100 respondents, while phone surveys used for polling and consumer research are often at 10% or below.

The initial target list for MaRS venture clients was extracted from the MaRS CRM database and verified by each relevant MaRS sector (ICT, Cleantech and Health). The list was in compliance with the Ministry of Research and Innovation policy, which outlined that target lists for the annual survey include (1) currently active clients, as well as (2) all new clients acquired in the fiscal year. Furthermore, the ventures were those that agreed to receive communications such as survey links, in accordance with Canada's anti-spam legislation (CASL). Finally, the ventures met the definition of "active ventures" as agreed on by working groups in which MaRS participated, such

as those created by the Innovation Data Partnership Ontario (IDPO). Regardless of the aforementioned restrictions around the creation of the list, it was robust and minimized the number of bounced emails.

ITEM NON-RESPONSE

The vast majority of questions in the survey were optional, meaning the respondent could leave the questions blank and proceed forward in the survey. As a result, some questions had high item non-response. For example, clients were asked their sales revenue for different territories (e.g. Canada vs. the World) and would enter an amount in one territory, but not the other. As a result, the territory which was left blank could be presumed to be zero.

Given the length and complexity of the survey, it was not feasible to make all questions required. Therefore, item non-response is expected, although it varies considerably based on the particular question.

FOLLOW-UP AND QUALITY ASSURANCE

A small number of outliers were filtered from the final results due to obvious nonsensical values. For example, one venture reported 240,000,000 in Federal Government funding. Since there are only a handful of companies in the entire country with this level of funding - none of which are MaRS ventures - such obviously erroneous data was removed.

In the case of survey data used for MaRS board reporting and analysis, various third party financial data was added for ventures that did not respond to the survey (Crunchbase, AngelList, CBInsights, CVCA). In other words, if a venture was on the target list, but did not respond, whatever private investment data was found for that venture for the relevant timeframe was added to the survey results. This only affected private investment and added 11% to the total capital. Every attempt was made to validate financial numbers with the clients themselves, or against available private sources of financial data

Investment Accelerator Fund (IAF) data was also added to the MaRS internal and board reporting, which affected jobs, capital and revenue numbers. Furthermore, in the case where a respondent answered the survey, but it conflicted with the IAF's own data, IAF's data would replace the original respondent data. In all, 100 IAF clients either supplemented or replaced MaRS survey data. All of IAF's clients that were supplied to Data Catalyst were used for this enhancement. Funded by the Province of Ontario, the IAF program is managed by MaRS and delivered through the Ontario Network of Entrepreneurs.

REPRESENTATIVENESS

Because certain variables do not have outliers due to the scales used, we consider them more robust estimates of the overall population of MaRS ventures. It is unlikely that a complete census of MaRS ventures would reveal significantly different proportions by gender, age, experience or nationality.

Because of the preponderance of accelerating ventures answering the survey, attitudinal questions may or may not be representative of the larger population. Venture ratings of the importance, satisfaction and impact of MaRS are consistently high, but it's possible that sampling bias influences these values higher than the overall population.

As discussed above, financial and employment metrics are not representative of the larger population, and should not be used for inference or extrapolation.