# Reimagining Survey Research: Transforming a traditional survey program through advanced analytics

Ryan Cristal • Survey Insights and Advanced Analytics • Office of Strategic Information, Research and Planning • Peace Corps
rcristal@peacecorps.gov

## Overview – Incorporating analytics at small organizations

**Leveraging advanced analytics in the context of an existing survey program:**

- **Requires building institutional support**
  - Speak to existing organizational mandates for:
    - utilizing evidence-based decision-making
    - growing the capacity to utilize data operationally
  - Strive for "quick wins" that inspire the confidence of stakeholders
    - Demonstrate concretely how the quality and utility of data improved
    - Deliver analyses that were not possible before, of issues of importance to the organization and key stakeholders
- **Requires data availability**
  - Analytics are robust if data and data sources (preferably multiple) are robust
- **Requires having the right expertise within the organization**
  - Analytics talent to do the work may be at a premium
  - Identifying experience on the stakeholder side will also drive success

## Results

**<u>Leveraging administrative and other data sources improves survey data quality…</u>**

- Raw data quality improvements

> *Example: Known response rate of demographic group X:*
> *Prior year: 73%  Implementation year: **<u>95%</u>***

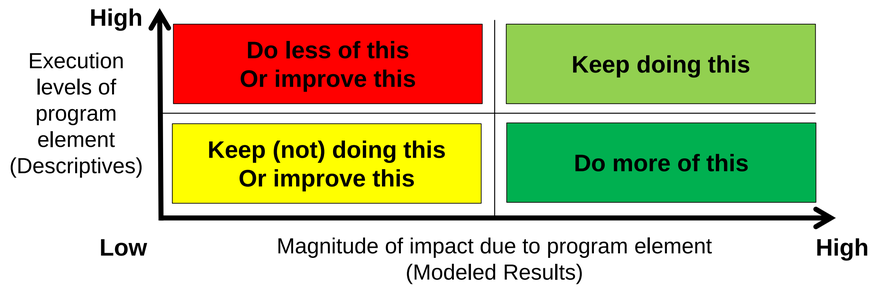
- Survey results production and validation improvements
  - Ability to employ multiple data sources enables additional options for weighting, imputation, non-response investigation, etc.

**<u>… and also improves survey data utility</u>**

- "Quick win" analyses of survey data in conjunction with other data sources
  - Often involved topics not previously analyzed or quantified
- Quantitative modeling of survey data to ascertain program effectiveness
  - Adapt techniques used in for-profit program effectiveness and ROI estimation
- Applying modeling techniques to identify relationships among data sources
  - Longer-term effort to assess data portfolio strengths and address whitespace

## Example: Quantitative modeling of survey data to ascertain program effectiveness

### Quantitative models can aid the decision-making process

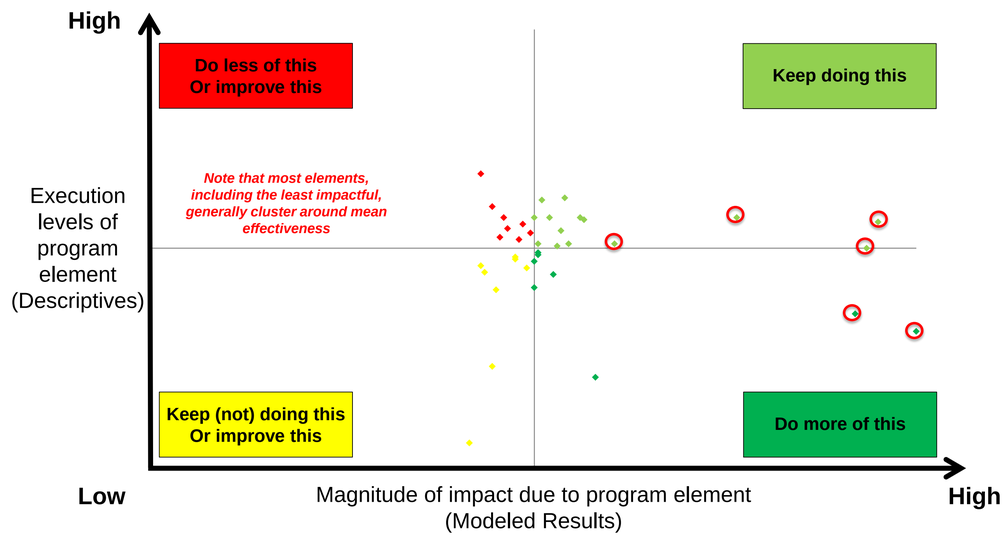



- *"Do more of what works, do less of what doesn't"* is a reasonable strategy for optimizing overall program performance. But, ***be careful!***
  - Models are aids to, not replacements for, judgement!
  - Optimizing performance can come from doing more of the strong elements ***<u>or</u>*** from improving the weak elements
  - Some elements are high-performing but not scalable (or they're expensive!), some elements are low-performing but are a "cost of doing business"
  - Balance short-term tactics and long-term strategy

### Uncovering insight from data to inform decisions

**Analyzing omnibus survey data utilizing multivariate regression modeling reveals that certain elements disproportionately drive overall program effectiveness**

***Optimize program effectiveness by focusing on the elements circled in red***




## Example: Predictive analytics in survey results production

### Applying predictive modeling to address survey data quality

***Situation:*** *Some responses to a survey question contingent on demographic classification are potentially impacted by distribution frame ambiguities*

***Solution:*** *Develop several predictive models utilizing differing analytic methods. Compare modeled predicted responses to each other and the actual survey responses. Formalize* ***conservative*** *decision rules to determine adjustments, based on model convergence*

**Example decision rules for determining survey response magnitude adjustments based on predictive modeling results**

| Logistic Regression Predicted Response | Decision Tree Predicted Response | Survey Actual Response | Adjust response? |
|---|---|---|---|
| 2 | 1 | 1 | No<br>At least one model agrees with actual, prioritize actual response |
| 1 | 3 | 2 | No<br>Models do not agree on directionality of magnitude |
| 0 | 0 | 1 | No<br>Models agree on directionality, but magnitude cannot be zero based on actual response |
| 2 | 3 | 1 | Yes<br>Models agree on directionality, increase actual by one step (i.e., "2") |

### Abstract

While the utilization of advanced analytics to inform decision-making is well established in the private sector and at larger public sector/non-profit organizations, the implementation of these capabilities is at a nascent stage in many smaller institutions. There are multiple reasons for this lag, relating to both technical capacity and organizational will.

Run by the USA government, the *Peace Corps* volunteer program has sent over 230 thousand Americans to 141 countries to provide technical assistance and foster intercultural understanding. While global in scope, the Peace Corps is a relatively small US government agency with limited resources devoted to data collection and institutional research. Despite these challenges, in a short time period, the Peace Corps has significantly improved the data quality and analytic utility of its survey insights program through the judicious application of data linkage and quantitative modeling techniques.

Transforming the Peace Corps survey program required careful cultivation of organizational support by demonstrating the value and promise of new techniques. This presentation will discuss the conditions that made this initiative more likely to succeed, the challenges faced, and the analyses produced. Prior to this, several necessary pieces were already in place: an institutional leadership with a strategic mandate to acculturate evidence-based decision making within the organization, an established large-N annual survey program, and multiple sources of administrative and program performance data.

Building upon this base, there were opportunities to begin socializing the organization to the possibilities of treating multiple data sources holistically and leveraging advanced quantitative approaches to decision-making analysis. One strategy was to utilize what was already available in order to generate new insights about the organization and actionable recommendations for improvement. In this case, multivariate regression modeling of large-N survey data was employed to investigate drivers of Peace Corps Volunteer effectiveness, an approach adapted from program effectiveness and return-on-investment analysis techniques more common in the for-profit context.

Concurrently, approval was secured for technical improvements to survey methodology, for example, implementing the use of administrative data capabilities to directly build and administer survey distribution frames. Care was made to proactively expose stakeholders to the immediate benefits. Improved data quality and analytic value were showcased, emphasizing organizational business issues not previously addressable in the absence of respondent-level data linkage of survey data with administrative, demographic, training, programmatic and other performance data.

The progress made in reinvigorating the Peace Corps' survey program has produced higher data quality, fresh insight, and greater stakeholder confidence, setting the stage for further capabilities growth. The presentation will close with current efforts to use modeling analytics to explain and predict survey non-response, to identify and address agency metric deficiencies, and to enable future analytic decision-making initiatives.