Methods Employed for Impact Monterey County's (IMC) Adult Experiences and Aspirations Survey and Online Data Availability

Questionnaire design, selection of participants, data collection and analysis

The IMC Data Team provided oversight and leadership in the creation of the survey questions, with input from the IMC Steering and IMC Engagement Committee members. Efforts were made to consider all sub-county geographic areas and diverse population groups, e.g., age, gender, ethnicity, household income, etc. in the distribution of the survey countywide.

The survey followed a convenience sampling approach and not a probability sampling method. The survey was open to all Monterey County residents in electronic and paper formats, and the IMC Engagement Committee led the task of promoting the survey across different geographic and demographic groups within the county. Survey data were collected between July 7, 2014 and October 31, 2014 through four survey instruments: two distributed online (in English and Spanish) and two paper versions (in English and Spanish).

An online survey application (Survey Monkey) was used to collect responses electronically. An initial email link was sent to United Way staff who then distributed it via email to members of the various IMC committees to forward to public agencies and nonprofit organizations and their constituents, family members, friends, colleagues, students, neighbors, etc. A survey link was also available on the United Way of Monterey County website (and other IMC Steering Committee agency/organization's websites?). Paper surveys were distributed at churches and libraries with the intent to extend availability to populations with limited internet and electronic access. These paper surveys were then collected by United Way volunteers and manually entered into the online Survey Monkey application by United Way staff. Multiple news media and social media outlets were utilized to promote survey participation.

The survey included closed-ended questions that provided respondents with options for a specific answer ("*pick the option that best describes your opinion*" or *"select all the options that apply from the list below")* and open-ended questions that included a comment box in which respondents could write unlimited responses in their own words. Many of the closed-ended items had an option for respondents to check "Other" and write in an answer. For the purposes of the Adult Experiences and Aspirations Report of May 2015, the "Other" information was not content analyzed for assignment to existing or additional categories of responses.[1] For the public release of the data via the Impact webpage, these results were analyzed and categorized as appropriate.

Respondents were required to provide their ZIP Codes which were used to construct regional clusters. For the written report, post-stratification weighting was used to balance representation by area of residence. Responses were assigned a weight based upon their geographic

---

[1] Please see Adult Survey Form for a complete reference to closed-ended questions with "Other" options.

(approximate supervisorial district) location, as well as Public Use Microdata Areas (PUMA) census data. [2]

The weight assigned to each response in a regional cluster was calculated by dividing the proportion of population residing in the respondent's geographic cluster by the proportion of survey questionnaires completed in the respondent's geographic cluster.

In the analysis, closed-ended item (quantitative) responses are expressed in percentages and calculated using the weighted dataset. All quantitative analyses were conducted using SVY commands in STATA version 13.1.[3]

Open-ended item (qualitative) responses, which were not part of a closed-ended item, were analyzed using NVIVO version 10, and themes for each item were extracted using a two-step process. First, word frequency analysis was used to obtain the top 3-4 most frequently occurring words in responses to an item (word clouds) and second, word-tree analysis was used to further explore the context in which each common word was used. The themes that resulted from the NVIVO analyses reside separately from the STATA-analyzed dataset.

Data Availability Online – Phase 2 of IMC

IMC's core values indicate that the survey data need to be useful to the local community, have detailed user-friendly summaries, and be available online. To this end, steps have been taken to develop online views of the data that utilize cross tabulations by demographics, residence, and population characteristics, that were not done in the published report that was county-wide in scope.

A Phase 2 Data Group was formed to carry out the steps required for this activity. Methods used to reorganize the data and transform them for ease of upload to an online system (Tableau[4]) included the following:

1) STATA files were converted to the format for the Statistical Package for the Social Sciences (SPSS)[5]. The data were reviewed and it was decided that the quantitative data (closed-ended items) would be used for the online system. During this review, it was determined that there were numerous responses of "Other" on closed-ended questions that were more accurately placed in one or more of the existing categories on the closed-ended answers. It was decided by the Data Group that these items be content analyzed to improve the accuracy and completeness of the data.[6] Content analysis was conducted

---

[2] The Carmel Valley ZIP (93924) falls in supervisorial district 5, with Carmel, Monterey and Big Sur, but it was assigned in the Adult Survey Report of May 2015 to the PUMA area for the south region of the county. For the public release of the data via the Impact webpage, these geographic areas were re-categorized as appropriate. For more information on the geographic clustering, contact Ignacio Navarro, Ph.D. CSUMB, Institute for Community Collaborative Studies.
[3] Copyright 1996-2015 StataCorp LP
[4] Copyright Tableau Software Incorporated and its Licensors. All rights reserved
[5] Copyright IBM Corporation 1994,2015
[6] Please see the Adult Survey form for the close-ended items with "Other" options.

using a dual review process – one reviewer assigning categories from the written answers, and a second reviewer assessing the accuracy of the first. Conflicting assessments were discussed and resolved to the satisfaction of both reviewers. The data were re-coded on the SPSS master file to reflect the content analysis results. For example, many respondents wrote in that they were spouses of someone with employer-based health insurance. These "other" answers were reassigned to the employer-based insurance category.

2) It was decided that the weight used in the county-wide report was not appropriate to employ in the production of detailed views, as the weights were only significant when looking at the whole county without the ability to filter by demographic characteristics.[7] The geographic areas were reviewed. Assignment of ZIP code 93924 (Carmel Valley) was made to the 5th area (Carmel, Monterey, Big Sur) for the purposes of the online system, as this was deemed to be a correction.[8] A second geographic variable ZIP Code, was constructed as follows. If the number of cases in the ZIP code was fewer than 50 and the ZIP code area was adjacent to an area with common demographics, that ZIP was combined with the adjacent ZIP code. This resulted in combining the following areas – Bradley and San Ardo with King City; Aromas with Castroville, Moss Landing with Pajaro, Chualar with Gonzales, Carmel-By-The-Sea with Pebble Beach, Big Sur with Carmel (93923).

3) In particular, the close-ended item tapping Race was examined for the frequency and content of the "Other" answers. It was found that several hundred respondents considered their Hispanic ethnicity as a race category. The resulting variable, Race, includes a category for "Hispanic".

4) The data were also reviewed for feasibility of restructuring to an online system. This was accomplished by using the Restructure function in SPSS. Briefly, this function converts a dataset that contains one record per individual (with columns representing answers), to a multiple-record per individual dataset with each row containing the response to an item. This process is termed normalizing[9], and enables survey data items to be loaded into visualization software such as Tableau.

5) A review of the proposed use of and confidentiality of the data was conducted - A Confidentiality Statement was written and approved. The following verbiage was finalized to be published with all materials pertaining to data requests and online use of the data.

---

[7] The user will be advised that there will be slight differences in counts and percentages when comparing the online data to the published report.

[8] The user will be advised that the counts and percentages for south County and for Carmel/Monterey/Big Sur will be different for the online data as compared to the report. There were 99 respondents from this ZIP.

[9] Wexler, S. Visualizing Survey Data, White Paper, 2014, http://www.tableau.com/sites/default/files/media/using_tableau_to_analyze_and_visualize_survey_data_clean_with_comment_sw_updatedjune13_2014_0.pdf

"The data shown from IMC are intended for the purpose of learning and continuous improvement around shared community goals. These data are a reflection on the entire county, which shares accountability for the current reality and for future results. They are not, and cannot be used as, evidence of the success or failure of one particular program, organization, or population and they should not be used in public or media communications without full and complete context, citation, and clear acknowledgement of the broadly shared accountability for results.

Please note: Data requests for analyses outside of the scope of the online system can also be submitted for consideration. Those requests for data in line with IMC goals and capacity will be processed. Attribution to the IMC Assessments will be expected and specific citation language will be provided. All data provided will be on the aggregate level with the criterion that cell counts of individual responses must be 15 or larger."

6) Adult Survey data including demographics, population characteristics, and survey responses were uploaded using Tableau 9.1