

Analysis Report

Data Transformation

The results are embedded in the Excel file delivered along with this report. The first step was to make the data ready to be analysed. Frequencies of responses were calculated for each question in the survey and whenever a particular answer option had been used by less than 5% of the total sample (21 responses or less), it was merged with other options to ensure a minimum representativeness of each response level. This procedure ensures that calculated numbers (counts and percentages) are not distorted by small sample sizes. One example of this data transformation is the 'Industry Type' question, where many sectors with very few respondents were merged into a 'Other' category and only those sectors with more than 21 responses were kept.

Whenever a question allowed more than one option to be marked ("Select all that apply"), this question was transformed on several binary variable (Yes/No) representing the choice of a particular category.

Results

The first four tabs of the Excel file show the crosstabulations. The first contains the percentage of respondents for each compared category (B2B, B2C and both or Company Size) that answered each survey option. The second presents the same numbers but with a color code to facilitate visualization. Numbers in red are below the average number for the categories being compared (either primary audience or company size). The third tab shows the percent numbers with their respective counts and the fourth shows the mean scores (averages) for all the questions that requested a numerical answer.

The tabs in Green show graphs based on the crosstabulations and the tabs in Blue shows the percentages and means supplemented with a statistical test, which is shown as subscript letters. Whenever a pair of numbers do not share an equal subscript letter for a given row, that means that the numbers for that pair of categories are statistically different. For instance, for the question "What percentage of your revenue comes from the following sources: - Inbound marketing" on the last tab, the mean percentage for B2B was 27.435 and for B2C it was 41.691. Since their subscript letters are not the same (a and b), that means that these percentages are statistically different. For the 'Both' category, the mean percentage is not statistically different from either B2C or B2B since a and b are also present as subscript letters. The same applies for the frequency of

responses on the tab 'Table_%_wStatistics': whenever the letters next to the 'Count' numbers are different, that means that the respective percentages are different.

SAMPLE REPORT - Rafael Data Analysis Portfolio