

City of Seattle Report on Information Technology Access & Adoption

INSTRUCTIONS FOR USING THE DATA SET

Do not pull down the data set and use it without using the code book. When you run the data set, be sure to refer to the Code Book for invalid or missing values for each question item (variables).

You must recognize that the data here is unweighted; to replicate the figures in the City's Information Technology Access and Adoption Report and get values that are representative of Seattle population, you should use the weights. There are 2 weighting variables, one that allows you to use the phone and online survey together (labeled "weightfin") and one that is for only the phone survey (labeled "wgtphoneonly", which excludes or zeros out everyone in the online survey).

Weighting

Weights were calculated with the aim of producing a dataset that reflects the population of Seattle in terms of ZIP code distribution, age, education, race/ethnicity, and income. In the initial step of this iterative process, Seattle population figures were gathered from City resources and deviations of the telephone survey sample from the City distribution were computed for each demographic. This was done by subtracting the category's prevalence in the sample from that category's prevalence in the population. For example, 1.64% of the City's population was estimated to live in ZIP code 98101, compared with 1.99% of the sample population. The difference between the survey value from the population value is -0.00365. For the purpose of weights, the average deviation of the categories was needed, so each difference was squared. For ZIP code 98101, the value squared was 0.00001264. The average of the squared ZIP code differences was 0.000055125. That average, which ranged from a low of 0.00055125 (ZIP code) to a high of .009732 (education for those under 25) provided an estimate of how closely the unweighted distribution of the sample matched the City's population.

Starting with the demographic with the greatest deviation, weights were calculated for each category that would reduce the sum of the squared deviations in the target demographic category to 0. For example, 33 people (4.41% of the respondents over 25 years of age) reported less than a high school education, compared with 7.70% of the City's population. Thus each of those 33 people would need to represent some of their neighbors as well as themselves. The weight that gives those 33 an appropriately "stronger" voice was 1.75. At the same time, 498 (66.58% of respondents over 25) reported at least a four year college degree, compared with 56.2% of the City's population. Thus the views of that group would be overrepresented in this sample without a correction. The weight that corrects this overrepresentation for this group is 0.844, somewhat "weakening" the voice of each respondent in this group. After applying these weights, the deviation from the population values for education was 0, but the differences between the sample and the population values had also shifted for each of the other demographics. The next step in this iterative process was to recalculate the average of the squared deviations for each demographic category after applying the weights for education. The

average of the squared deviations ranged from 0.0 (education) to 0.00410 (age). Using the method described above for education, weights were calculated for the different age categories. These weights were combined with the weights for education by multiplication and applied to the sample.

This process was applied next to the income categories and finally to the race/ethnicity categories until the average deviation of any category was within the 95% margin of error for that demographic. The demographic with the largest average deviation following the weighting procedure was education among those older than 25. Table 1 provides the education distribution of the population over 25, and the corresponding distribution of the sample.

Table 1. Population and weighted sample distribution of education for those 25 and older		
Education Category	Population	Sample with final weights
Less than high school	7.7%	6.3%
High school graduate/ GED	11.7%	10.8%
Some college or a two-year degree	24.4%	24.6%
At least a four year degree	56.2%	58.0%

This process was repeated for the online survey so that each dataset could be analyzed independently or together with appropriate distribution.