

## Logan County Needs Assessment - Sampling Plan

### Brief Description of Logan County/Needs Assessment Project

Logan County is located in northwest Ohio just north of Champaign County and west of Union County. The county is just less than 500 square miles, extending north to Indian Lake and south to West Liberty. The county seat is Bellefontaine, with a population of approximately 13,000 people.

The current project is a combined effort of a number of county and community agencies to administer a county-wide needs assessment in the areas of preventative medicine, general health care, mental health, drug and alcohol use, physical safety, utilization of existing community resources, and more. The survey is a combined effort in that results from the survey will be used to further the needs of a number of partner agencies. Being a “needs assessment”, results from the survey will be used to establish trends over time, provide data to inform direction of service provision, and provide information on the current status of services in the county.

### Definition of “Stratified Random Sample”

The survey will be administered to a stratified random sample. A stratified random sample is a sample where respondents are chosen randomly, but the random assignment is not done with the sample as a whole, but within subsets of the sample – in this case census tracts. Basically, within each census tract a similar percentage of the population will be chosen to participate. After initial survey results are obtained, checks will also be made to ensure that the sample is representative, both within census tracts and as a whole, on the basic demographic characteristics of gender, race, and income level. If discrepancies are found, an attempt will be made to gather additional survey information from underrepresented groups. If survey results are still not representative, individual data will be weighted as needed so that results reflect a sample representative of 2010 census data.

### Census Tracts

There are a total of 11 census tracts in Logan County, corresponding to census tracts 9838 to 9848 respectively. The following shows the estimated population of each tract via the 2010 Census.

Tract Number	Geographic Area	2010 Population
9838	Rushsylvania and West Mansfield	3,862
9839	Huntsville and Belle Center	4,561
9840	Indian Lake State Park	2,536
9841	Lakeview and Chippewa Park	2,813
9842	Russells Point	1,898
9843	Degraff, Quincy, and Lewistown	5,849
9844	Bellefontaine (west)	5,117
9845	Bellefontaine (central)	4,343
9846	Bellefontaine (east)	5,086
9847	Zanesfield, Valley Hi, and East Liberty	5,355
9848	West Liberty	4,438
	<b>Total Population</b>	<b>45,858</b>

There are a number of options of the population percentage to sample within each census tract. For the purposes of this project, three options will be considered – a 1%, 2%, and 5% sample respectively. If budgets allow, a larger sample could be considered. These options seemed realistic, however, within time and budget constraints.

**Option A – 1% Sample**

Option A would sample approximately 460 subjects as follows:

Tract Number	Geographic Area	Sample Size	Confidence Interval
9838	Rushsylvania and West Mansfield	39	+/- 13.6%
9839	Huntsville and Belle Center	46	+/- 12.5%
9840	Indian Lake State Park	25	+/- 17.0%
9841	Lakeview and Chippewa Park	28	+/- 16.0%
9842	Russells Point	19	+/- 19.5%
9843	Degraff, Quincy, and Lewistown	58	+/- 11.1%
9844	Bellefontaine (west)	51	+/- 11.9%
9845	Bellefontaine (central)	43	+/- 12.9%
9846	Bellefontaine (east)	51	+/-11.9%
9847	Zanesfield, Valley Hi, and East Liberty	54	+/- 11.5%
9848	West Liberty	44	+/- 12.8%
	<b>Total Sample</b>	<b>459</b>	<b>+/- 4.0%</b>

Confidence intervals are a way of estimating margin of error. In political polls, often results are posted with a “margin of error” of a certain number of percentage points. For instance, one might predict that an issue would win by a vote of 58% for and 42% against with a margin of 4 percentage points. That “margin” is a confidence interval. Typically, the interval is quoted with a 95% certainty.

Confidence intervals were estimated using a normal approximation interval on a binomial (e.g., yes/no) variable, multiplying the z-value at the 95% confidence level with the standard error. The middle proportion values of .75 and .25 were used. For example, for a question such as “Were you ever diagnosed with high blood pressure?”, if 75% of respondents in census tract 48 (West Liberty) answered yes, one can reasonably assume that the actual percentage of people with high blood pressure in the census tract is between 62% and 88%. Responses with less or more agreement will lead to different confidence intervals.

The formula for confidence interval is as follows:

$$\begin{matrix} \text{---} \\ \text{---} \end{matrix}$$

$P_s$  is the probability of success in the sample (“yes” response) and  $Q_s$  is the probability of failure in the sample (“no” response) – basically .75 and .25. The size of the sample is signified by  $n$ . So, the formula for confidence interval is:

$$\frac{P_s - Q_s}{\sqrt{\frac{P_s Q_s}{n}}}$$

For example, the confidence interval of the total sample would be 1.96 times the square root of .1875/459, or .040 (+/- 4.0 percentage points). As one can see, the overall confidence interval is within acceptable limits. However, with small samples, the confidence interval becomes extremely large, making estimates within some of the smaller census tracts problematic.

**Option B – 2% Sample**

Option B would sample approximately 920 subjects.

Tract Number	Geographic Area	Sample Size	Confidence Interval
9838	Rushsylvania and West Mansfield	77	+/- 9.7%
9839	Huntsville and Belle Center	91	+/- 8.9%
9840	Indian Lake State Park	51	+/- 11.9%
9841	Lakeview and Chippewa Park	56	+/- 11.3%
9842	Russells Point	38	+/- 13.8%
9843	Degraff, Quincy, and Lewistown	117	+/- 7.8%
9844	Bellefontaine (west)	102	+/- 8.4%
9845	Bellefontaine (central)	87	+/- 9.1%
9846	Bellefontaine (east)	102	+/- 8.4%
9847	Zanesfield, Valley Hi, and East Liberty	107	+/- 8.2%
9848	West Liberty	89	+/- 9.0%
	<b>Total Sample</b>	<b>917</b>	<b>+/- 2.8%</b>

With this second option, most of the confidence intervals are all under 10% (the exceptions being the smallest census tracts). In addition, the confidence interval of the entire sample is very low – 2.8%.

### Option C – 5% Sample

Option C is the most extensive sample, and would include approximately 2,300 subjects.

Tract Number	Geographic Area	Sample Size	Confidence Interval
9838	Rushsylvania and West Mansfield	193	+/- 6.1%
9839	Huntsville and Belle Center	228	+/- 5.6%
9840	Indian Lake State Park	127	+/- 7.5%
9841	Lakeview and Chippewa Park	141	+/- 7.1%
9842	Russells Point	95	+/- 8.7%
9843	Degraff, Quincy, and Lewistown	292	+/- 5.0%
9844	Bellefontaine (west)	256	+/- 5.3%
9845	Bellefontaine (central)	217	+/- 5.8%
9846	Bellefontaine (east)	254	+/- 5.3%
9847	Zanesfield, Valley Hi, and East Liberty	268	+/- 5.3%
9848	West Liberty	222	+/- 5.7%
	<b>Total Sample</b>	<b>2,293</b>	<b>+/- 1.8%</b>

With a larger sample comes a smaller confidence interval. This is a huge advantage. It allows for a much greater amount of accuracy in estimating population responses. However, a large disadvantage is cost. The survey is rather large (28 pages) and the cost of printing and mailing approximately 2,300 surveys may, in fact, be prohibitive.

### Overall Conclusion/Recommendation

Because of the large confidence intervals in a number of the census tracts in Option A, Option A is not recommended. There is no statistical rule against it, it would just lead to a large amount of uncertainty in population estimates. If budgets allow, Option C is would provide the best estimates of population needs. However, Option B is a satisfactory alternative.