Exercise 2: Planning and Analyzing Two Samples

Item 1. Planning Your Sample, Based on Number of Inspections You Can Do in Each Round.

Hampsachusetts, a participating State, has a population of 500 facilities. Use the Sample Planner to estimate the margin of error for differences in proportions (at 95% confidence level) associated with each of these numbers of inspections:

Number of Inspections (Each Round)	Max. Margin of Error For the Difference	Max. Width of Confidence Interval for the Difference (i.e., MoE X 2)
30	+/- 25.0%	50.0%
50	+/- 18.8%	37.6%
100	+/- 12.5%	25.0%
150	+/- 9.5%	19.0%
200	+/- 7.6%	15.2%

Anything interesting about these results? Is there a point at which the confidence interval is too large to be of practical use? The usefulness of each margin of error or confidence interval will be determined by your own purposes. Perhaps think about how wide is too wide for whatever decisions you might need to make. You can also think about the margin of error for the difference as the minimum difference you will likely be able to detect. E.g., with 30 inspections in each round, even if you observed a difference of 25% between the two samples, you may not be able to say that that difference is significant (i.e., that a difference in the population actually occurred).

Item 2. Planning Your Sample, Based on Desired Margin of Error.

Connectimont, a participating State, has a population of 500 facilities. Use the Sample Planner to:

1. Identify the number of inspections required for achieving a margin of error of +/- 7.5%, at 95% confidence level.
Answer: 203 inspections

2. Connectimont decides that's too many inspections. They'd like to consider a 90% confidence level, same margin of error. How many inspections do they need?
Answer: 163 inspections.

3. Taking the answer from #2, what's the margin of error they would get with a 95% confidence level?
Answer: +/- 8.9%

Anything interesting about these results? If they had to choose between #2 and #3, what would you recommend? Interesting: It's much easier to achieve desired margins of error with lower confidence levels. The question is whether these confidence levels are sufficient for your purposes.

Item 3. *Comparing Two Samples from a Single State, Across Time.*

Using the information in the gray part of the table, use the Results Analyzer to calculate results for three different indicators for Hampsachusetts, then report what's interesting about the results. Assume a population of 500 facilities in both rounds, and a confidence level of 95%. Hint: think about what you calculated in Item 1 and 2, and think about what's happening to the half-width of the confidence interval here below.

State	Indicator #	Overall Sample Size (Each Round)	Effective Sample Size[1] (Each Round)	Round 1: # Meeting Criterion (Positive Hits)	Round 2: # Meeting Criterion (Positive Hits)	% of Positive Hits in Round 1	% of Positive Hits in Round 2	Observed Difference	Margin of Error	Conf. Interval (Lower Bound)	Conf. Interval (Upper Bound)	Different?	What's Interesting About This Result?
Hampsachusetts	1	100	100	50	60	50%	60%	10%	+/-12.3%	- 2.3%	+ 22.3%	NO	MoE slightly smaller than planned; planning tool makes most conservative assumptions about observed proportion.
	2	100	100	10	20	10%	20%	10%	+/-8.8%	1.2%	18.8%	YES	MoE decreased for same sample size and observed difference; proportions farther away from 50%.
	3	100	50	5	10	10%	20%	10%	+/-13.3%	- 3.3%	23.3%	NO	MoE increases substantially relative to indicator 2 (and is more than planned for) because of decrease in effective sample size (despite proportions close to zero).

---

[1] "Effective sample size" reflects the number of facilities in the sample for which the indicator is relevant. For instance, if you sample 100 facilities and the indicator is only applicable to 75 of those, your effective sample size is 75.

Item 4: Comparing Single Samples from Different States.

Using the Results Analyzer, compare results for imaginary indicator #4 between different states, and consider what can be learned from this exercise. How much does the sample planning in one state affect the ability of all states to draw comparisons? Do we need to plan for samples differently than if we are just planning to compare two samples over time within one state?

Data to Input into Results Analyzer

State	Pop. Size (Round 1)	Conf. Level	Indicator #	Overall Sample Size (Round 1)	Effective Sample Size[2] (Round 1)	Round 1: # Meeting Criterion (Positive Hits)	% of Positive Hits in Round 1
Hampsachusetts	500	95%	4	100	100	50	50%
Connectimont	500	95%	4	50	50	30	60%
Vermaine	500	95%	4	200	200	80	40%
Jerserado	1000	95%	4	100	100	80	80%
Rhodifornia	500	95%	4	30	30	18	60%

Answer Worksheet

State	Compare to State:	Observed Difference	Width Of Confidence Interval	Conf. Interval (Lower Bound)	Conf. Interval (Upper Bound)	Different?	What's Interesting About This Result?
Hampsachusetts	Connectimont	10%	+/- 15.7%	- 5.7%	+ 25.7%	NO	Margin Of Error is larger than in Sample Planner for Item 1 because of lower sample size for Connectimont.
Hampsachusetts	Vermaine	10%	+/- 10.3%	- 0.3%	+ 20.3%	NO	Even Vermaine having twice the sample size as Hampsachusetts does not allow detection of a 10% change with these proportions.
Connectimont	Vermaine	20%	+/-14.0%	6.0%	34.0%	YES	Change is detectable. Is it useful?
Jerserado	Vermaine	40%	+/-9.1%	30.9%	49.1%	YES	Change is detectable. Is it useful?
Jerserado	Rhodifornia	20%	+/-18.8%	1.2%	38.8%	YES	Change is detectable. Is it useful?

Note: In the interests of time, this exercise focuses on comparisons that illustrate points somewhat distinctly from comparisons that have been left off.

---

[2] "Effective sample size" reflects the number of facilities in the sample for which the indicator is relevant. For instance, if you sample 100 facilities and the indicator is only applicable to 75 of those, your effective sample size is 75.