HECA Sampling Strategy Summary

Sampling is a statistical method used to gain meaningful insights about a large population by analyzing a smaller subset of data. For example, when we collect a large and diverse enough sample of HECA data from an organization, we can make conclusions that apply broadly to the entire organization.

How many HECA opportunities do I have?

To determine the required sample size:

- **Start** by estimating the <u>total population size</u>, represented by the total number of HECA opportunities.
- **Consider** that each HECA is recorded for a crew performing a task.
- **Calculate** the total HECA opportunities by multiplying the number of worker-hours in a month by the average crew size and the average number of times a worker accumulates hours in a month.
- Based on feedback from the **EEI community**, the average worker accumulates 200 worker-hours per month and the average crew has 3 individuals.
- **Use** the following equation for estimating total number of HECA opportunities.

$$Number\ of\ HECA\ opportunities\ per\ month = \frac{Total\ number\ of\ field\ worker\ hours\ per\ month}{200\ hours\ per\ employee\ per\ month\ x\ 3\ employees\ per\ crew}$$

For example, if a business unit accumulated 300,000 worker hours in a month, we can use the equation above to estimate that there were 500 HECA opportunities that month.

How much HECA data do I need to have a representative sample

Once the population size, (i.e., number of HECA opportunities) is known, well-established statistical sampling formulas can be applied to estimate the required number of HECA entries needed. By choosing a confidence interval and making typical statistical assumptions (i.e., a 5% standard error), the sample size can be computed. Note that the confidence interval is the likelihood that the trends observed in the sample represent the actual trends in the entire organization.

To simplify the process of determining the required sample size for HECA, the table below was created. Since most EEI companies fit neatly into one of the following *monthly* worker hour (WH) segments, the associated HECA sample can be summarized for each segment.

WH Segment / per month	Confidence Level					
	95%		85%		75%	
	per month	per day	per month	per day	per month	per day
<50k	43	2	38	2	34	1
50k-250k	145	5	108	4	83	3
250k-500k	299	10	179	6	120	4

For example, if a business unit accumulates 300,000 worker hours per month and would like to measure their HECA score with 75% confidence, a minimum of 120 HECA entries must be collected per month accumulated by an average of 4 HECA entries per day.

Is my HECA data representative?

The extent to which HECA data reflect the work depends on the quality not only the quantity of data. The following best practices help to ensure a high-quality sample:

- **Calibrated assessors:** It is crucial that the those measuring HECA are adequately trained to follow the HECA rulebook and yield valid and reliable assessments.
- **Randomization:** Samples should be chosen without bias, ensuring each crew has an equal chance of being selected for HECA in a given work period.
- **Representativeness:** The sample should include some representation of the different types of work performed by the business unit (e.g., tasks, locations, work environments, etc.)
- **Consistent cadence:** Consistent and reliable measurements with a set cadence are essential for valid results.

Other Notes

1. If a business unit accumulates more hours than 500,000 worker hours, it should be broken down to smaller business units or geographical regions.
2. HECA is meant to apply to crews working in the field (e.g., warehouse, grid ops, customer service, etc.). If an organization cannot estimate the number of field worker hours, the organization may use the total number instead which will result in a conservative sample estimate.
3. HECA can be collected by safety professionals, third party observers, or field supervisors if they are trained and calibrated to collect consistent data. To avoid conflicts of interest, crew leads should not assess the crews that they directly supervise.
4. **The EEI Safety Measurement Team members committed to HECA data collection at minimum 75% confidence level for 1 Business Unit by the end of 2024.**