Statistical Habits of Mind

Hollylynne Lee & Dung Tran
Friday Institute for Educational Innovation
NC State University

A habit of mind is developed when a person approaches situations in similar ways so that a more general heuristic is accumulated over time. Some specific habits of mind are productive for engaging in while doing statistics. Thus, as both a learner of statistics and a teacher of statistics, we need to be developing the following habits for ourselves and our students:

- Always consider the context of data
- Ensure the best measure of an attribute of interest
- Anticipate, look for, and describe variation
- Attend to sampling issues
- Embrace uncertainty, but build confidence in interpretations
- Use several visual and numerical representations to make sense of data
- Be a skeptic throughout an investigation

The paragraphs below provide more detail about these statistical habits of mind and then map the habits into how they can be helpful in the four phases of a statistical investigation.

Context

Data are not merely numbers, but numbers in context. A data value of 33 has different meaning if the context of data is the age of mothers with their first child (likely considered a typical value) or the number of points in a basketball game scored by a single player (an unlikely value and may be a new record). Therefore, students should consider the role of context in every phase of a statistical investigation. Context is crucial for framing questions of interest and informs data collection. Context sheds light on data analysis and can be useful when deciding if a graphical display (and its scale) or numerical statistics make sense, or if any values should be considered as outliers and how they may distort the story. In the interpretation phase, claims should refer back to the context to make clear how a decision applies in that context and to help assess the reasonableness of any inferences being made.

Measurement

Good data is key to answering statistical questions. Without collecting appropriate data, any analysis may not have meaning and claims made could be based on inappropriate or invalid measures. When collecting data, the measurement should stem from contextual problems -- how to best capture the phenomenon of interest with the measurement. For example, students' academic score in a subject might be a proxy for students' learning, but not all of their learning and understanding about that subject. In addition, the use of appropriate tools to collect data such as the way to pose questions in a survey is crucial for good data. One might consider if different people making such measurements obtain very similar results and if the measurement is repeated, would it lead to similar data. In some cases, data have already been collected and statistical questions of interest stem from the investigation of those data. In such a case it is important to understand what was being measured and how the data were collected. Students should refer back the context when considering outliers and missing data in the analysis and interpretation.

Lee, H. S., & Tran, D. (2015). Statistical habits of mind. In *Teaching statistics through data investigations MOOC-Ed*, Friday Institute for Educational Innovation: NC State University, Raleigh, NC.
Retrieved from https://fi-courses.s3.amazonaws.com/tsdi/unit_2/Essentials/Habitsofmind.pdf

Variability, Uncertainty, and Sampling

Doing statistics means making decisions under uncertainty. If there was no variation in the world, we would not need statistics. However, because things vary (e.g., mothers do not all have their first child at a certain age, and basketball players scoring performance varies from game to game) statistics can help us investigate questions about contexts in which we anticipate variability (e.g., Over the last few years, how do the highest paid basketball players fare in their points scored per game?). Sources of variation could occur naturally in the context; however variation can also stem from measurement (e.g., accuracy of a measurer or measurement device), sampling, or by data entry errors. Statistical investigations seek to explain, predict, and control variation inherent in data. Well-designed sampling helps avoid self-selected or biased samples. Thus, the use of random sampling is important, as well as consideration of the size of a sample. The goal of well-designed study is to explain that something happens just by chance, not by errors. When making final claims or inferences, students should be aware of the variation inherent in data, and acknowledge that any decision made from data is under uncertainty. We just can't say for sure that something is certain and true.

Visuals and Trends

One important aspect when analyzing data is creating meaningful graphical representations and numerical statistics to examine the structure of data. It also involves comparing, connecting, changing, and transforming data representations to highlight informative trends or patterns. Different representations can shed light on different aspects of the data. For example, using boxplots could display a five number summary, overall variation, and possible outliers. However, a dot plot or histogram may be better for displaying how data is clumped and spread out throughout the entire range. Students should be encouraged and afforded opportunities to explore their data using multiple graphical representations, determine which are the most appropriate for the data, and observe overall trends. Graphical visualizations should precede computing numerical statistics. However, deep analysis should use both visual displays and numerical statistical computations to describe trends and patterns in data.

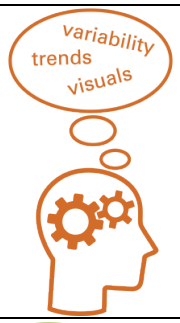
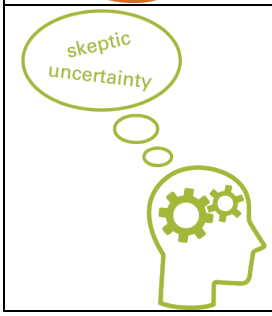
Skeptic

Being a skeptic means constantly looking for flaws in a statistical investigation. Skepticism is especially critical in making claims and interpretations based on data, and judging those made by others, particularly those in media reports. Students should be nudged to be skeptical about claims -- to worry about motivation, predispositions, and objectiveness of the writer, which might affect the trustworthiness of the claims. Students should ask themselves: "To what extent do I trust this information?" This question could be addressed by examining the choice of measure in relation to its context, the design of a study, sampling techniques used, and the analysis of data.

Using Habits of Mind in Phases of an Investigation

The following table gives a brief reminder of how the statistical habits of mind can be useful in the four phases of a statistical investigation.

Lee, H. S., & Tran, D. (2015). Statistical habits of mind. In *Teaching statistics through data investigations MOOC-Ed*, Friday Institute for Educational Innovation: NC State University, Raleigh, NC.
Retrieved from https://fi-courses.s3.amazonaws.com/tsdi/unit_2/Essentials/Habitsofmind.pdf

	***Pose Questions*** • **Context:** Ask contextually-based questions that call for the use of data to answer. • **Variability:** Seek to explain and control variability.
	***Collect Data*** • **Measurement:** Consider how to best measure attributes in a context for answering a question. • **Measurement:** Use appropriate tools (physical and online) to collect and manage data. • **Sampling:** Consider sample size – it matters. • **Sampling:** Use random sampling to help control bias. • **Sampling:** Identify and account for sources of potential variability in data collection methods.
	***Analyze Data*** • **Visuals:** Use appropriate tools strategically for creating multiple representations. • **Variability:** Coordinate graphs and statistical computations to reason about distributions in the aggregate. • **Trends:** Look for patterns and relationships within and among variables. • **Context:** Consider context of your question to identify measurement issues (missing data, outliers).
	***Interpret Results*** • **Context:** Reason quantitatively and make arguments supported by data. • **Context:** Make a claim connected to the context of the questions. • **Uncertainty:** Account for uncertainty in a claim (be confident but not certain). • **Skeptic:** Check the reasonableness of a claim (skepticism)

Lee, H. S., & Tran, D. (2015). Statistical habits of mind. In *Teaching statistics through data investigations MOOC-Ed*, Friday Institute for Educational Innovation: NC State University, Raleigh, NC.
Retrieved from https://fi-courses.s3.amazonaws.com/tsdi/unit_2/Essentials/Habitsofmind.pdf

References

Ben Zvi, D., & Garfield, J. (2004). Statistical literacy, reasoning, and thinking: goals, definitions and challenges. In D. Ben-Zvi & J. Garfield (Ed.). *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 3-15). Kluwer.

Burrill G., & Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.), *Teaching Statistics in School Mathematics - Challenges for Teaching and Teacher Education* (pp. 57-69). Springer Netherlands.

Cuoco, A., Paul Goldenberg, E., & Mark, J. (1996). Habits of mind: An organizing principle for mathematics curricula. *The Journal of Mathematical Behavior, 15*(4), 375-402.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). Guidelines for assessment and instruction in statistics education (GAISE) report. *American Statistical Association*. Online at http://www.amstat.org/education/gaise/.

Shaughnessy, M., Chance, B. L., & Kranendonk, H. (2009). *Focus in high school mathematics: Reasoning and sense making in statistics and probability*. National Council of Teachers of Mathematics.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223-248.