TECH TIP

Data Science Education Principles

Philosophy

The Tech's approach to data science education is an exploratory and interdisciplinary one. Like our mission to inspire the innovator in everyone, we want to ensure that youth engage in hands-on, open-ended solutions driven by data.

Data science involves humans and computers collecting, processing, analyzing, and utilizing data (data science skills) to understand and solve problems. We do this by using computational thinking (CT) skills to make sense of data. This will help spark learners' curiosity and inspire them to learn more or engage in action. We focus on real-world problems and questions, and use data to surface potential approaches to answer complex questions and address inequities. Through this process students are exposed to different careers and discover the role of data science in every field.

Principle I: Data is everywhere

Data is present in everything that we do — from taking a picture with friends to finding directions to your next destination, or deciding when it's best to do your grocery shopping. It's important to acknowledge that we live in a data-driven world, but learners can also use it as a tool to learn more about things they are interested in or issues that affect them. As educators we have the opportunity to incorporate data science into every subject or topic. In a social science class students may examine demographic and political data. They may conduct statistical analysis of weather patterns to understand climate science. Even the arts have a role as students create and develop data visualizations that accurately communicate information. When data science is authentically embedded into different subjects, it invites learners with lower confidence in their STEM skills to engage in these skills beyond their statistics courses.

Principle II: Call out connections to computational thinking, computer science, and artificial intelligence

The fields of data science and computer science are large and interconnected. Data scientists often use computational thinking skills such as pattern recognition and abstraction to analyze and visualize data respectively. When we explicitly make these connections, youth realize that they can develop proficiency in skills that can be used to problem-solve. This "call out" normalizes the academic language and expands access to data science, a sometimes intimidating field, by reducing the "fear" of the unknown.

Calling attention to the skills students are building can expand student understanding of the opportunities available to them. Computer scientists may need to create artificial intelligence models which require vast amounts of data, and yet computational thinking skills are necessary as models begin to create the algorithm it will be based on.

This explicit call out and connection demystifies complex concepts, reduces intimidation, and encourages students to see themselves as learners of these concepts. Ultimately this can broaden participation and foster a deeper understanding of the interconnectedness of modern technological disciplines. Youth who otherwise may never have considered computer science or data science may begin to see themselves in these careers.

Principle III: Acknowledge that data (in it's collection or preservation) can be biased due to human factors

Curiosity is a core tenant of our approach. Begin by asking students what they notice and wonder about data. Encourage them to think beyond the data provided to the data collection process and data visualization creation. Data can help tell a story and draw people in, but it might also have gaps — it's important to be conscious of the perspective that is being used. Questions such as who, what, when, where, why and how can help provide context to any data analysis. Acknowledge the limitations of data and embrace its complexity. Encourage critical thinking and ethical considerations of data in order to surface inequalities instead of perpetuating the systems that enforce them. This will empower students to see data as a tool they can use to take on systemic issues in their communities.

Principle IV: Focus on data science concepts not the tool

There are a variety of different types of data science tools. Familiarizing students with each of them takes significant time and may not benefit them in the long run with the pace at which technology is evolving. Instead, prioritize providing learners with opportunities to practice data science skills they need in order to analyze data and present it in meaningful ways that can spark action/curiosity. While data collection is a key part of the process, collecting and cleaning larger datasets can also be extremely time consuming and limit the connection to real world applications. Instead, The Tech prefers to uplift community-based data sets to showcase the work of local organizations and how they are tackling local problems.

Principle V: Take data-based actions to address real world challenges and inequities

Data can inform change and empower problem-solving. Students are often motivated to move beyond data collection and analysis to actions and solutions. Embrace this opportunity to engage them in issues that matter to them and their community. Grounding skills in data science within these real-world problems can increase participation and inspire students who might not typically be interested in STEM careers.

Principle VI: Value diverse perspectives and foster continuous learning and improvement

Provide students opportunities to share ideas in their own words, building confidence through both formal and informal communication. This approach to storytelling with data engages students to have a greater STEM selfefficacy. Create a safe space for open-ended discussions and analysis so that students feel comfortable bridging the gap between colloquial and academic language. Allow for multiple interpretations and perspectives — everyone needs access to data science!

For more resources and lessons on data science see **thetech.org/data-science-lessons**.