



## STUDENTS CREATE INFORMATION IN THE NEW DATA SCIENCE MINOR

The world is swimming in big data. The problem is, not many people know what to do with it.

Cal Poly is filling that gap with a new cross-disciplinary minor in data science. Students who graduate with the minor will be among the few who can help companies make sense of big data. The program is a collaboration between the Statistics and Computer Science departments and is tailored to meet the needs of this growing job sector.

### IT'S A TRULY INTERDISCIPLINARY, CROSS-COLLEGE PROGRAM, WHICH IS RARE.

"Graduates of the minor will really be among the first generation of formally trained data scientists," said Andrew Schaffner, a statistics professor who helped to develop the program.

Data science lies at the intersection of computing and statistics, and data scientists turn all forms of data — for example, numbers, text, sound or images — into usable information. Though data is constantly being produced — through internet searches, consumer transactions, bioinformatics and more — it can't tell us anything until it's been processed.

"What's missing is humans to process this data," said Alexander Dekhtyar, a computer science professor involved

in creating the program. "There aren't many people out there who have this breadth of skills."

The minor aims to give computing and statistics majors that broad skill set by enhancing the traditional programs. Statistics majors will add programming and database classes, while computer science majors will study probability and statistical modelling.

"It's a truly interdisciplinary, cross-college program, which is rare," Schaffner said. "The upper-level courses will be team-taught, giving students access to expertise in both fields in one class."

In true Learn by Doing style, students in the minor will work with real-world data throughout the program. During the minor's capstone experience, students will help actual customers understand large data sets.

Graduates will be able to work at any point in the data processing pipeline: collecting data, converting it into a usable form, visualizing and describing it, and finally interpreting it to help make data-based decisions.

"The demand for graduates with this skill set is only going to grow with the Internet of Things," Dekhtyar said. "When your car engine, your refrigerator and your house are all producing data, companies across all industries will need data scientists."

Cal Poly's first official data science alumni will enter the job market at the end of this school year with two years of classes under their belts. The minor is currently open to statistics, computer science and software engineering majors, with more majors to be added as the program grows. //

*Pictured: Statistics professor Andrew Schaffner works with data science students. Photo credit: Tenney Rizzo*