CAL POLY HUMBOLDT
University Senate

Resolution to Recommend Data Cleaning and Visualization Certificate
X-24/25-ICC- - February 11, 2025 – Curriculum Reading

RESOLVED: That the University Senate of Cal Poly Humboldt recommends to the Provost that the Data Cleaning and Visualization Certificate detailed in proposal 24-2462 be approved.

RATIONALE: Data is ubiquitous, and the field of data science has applications in nearly every other domain. Developing data cleaning and visualization skills is a way to set our graduates apart in a competitive workforce and build their technical portfolio. As a specific example, the field of biology has transformed where many experiments require analyzing and interpreting large biological datasets (e.g., genomics, metabolomics, etc.). These datasets can be messy and need to be cleaned before a meaningful analysis can be conducted. Visualizing data is helpful for extracting insights and clearly communicating results to stakeholders.

Humboldt has launched a new Data Science major but currently does not have a way to target students in other majors who would like to obtain marketable data science skills other than offering our courses as electives with no coherent program. Moreover, it is challenging to communicate to individual students which specific courses would be most helpful for specific career skills (this is usually done between students and advisors and requires the advisors to know the course specific content, the current field which is rapidly evolving, and what our department offers).

By selecting coursework and packaging it into a tangible certificate that students can list on a resume, we will better be able to serve our students. Additionally, our data science courses weave the theme of "Data for Good" throughout– this is directly in line with Humboldt's commitment to environmental and social justice and is good both for the University and for our students.

Certificate Description:
The Cal Poly Humboldt Data Cleaning and Visualization Certificate provides foundational concepts in computational thinking, computer programming, data cleaning and data visualization which are transferable and relevant across disciplines and industries. Co-requisite: MATH 101.

Includes the Following Courses:
- DATA 111 - Introduction to Programming and Computational Thinking for Data Science (4 units): Intertwines three perspectives: inferential thinking, computational thinking, and real-world relevance. Explore real data to understand relationships and patterns while learning critical skills in computer programming and basic statistical inference.

- <u>CS 111 - Computer Science Foundations 1 (4 units):</u> Introductory programming covering problem decomposition, control structures, simple data structures, testing, and documentation. Students design and implement a number of programs.
- <u>DATA 271 - Data Wrangling and Visualization (4 units):</u> Provides an intensive, hands-on introduction to data wrangling, cleaning and visualization using a contemporary programming language. Learn the fundamental skills required to acquire, transform, manipulate, and visualize data in a computing environment that fosters reproducibility. The overall goal is to create actionable data from raw sources and then perform exploratory analysis. These steps will include importing data, viewing data diagnostically, identifying outliers, imputing data, cleaning data, calculating basic statistics and creating informative plots.

Humboldt currently offers the courses we are suggesting in each certificate. They are all courses in the Mathematics Department or the Computer Science Department. There is currently excess capacity in these courses, so this is an efficient way to use existing courses.

Related Certificates:
- [Database and Data Analysis Certificate - 23-2474](#)
- [Machine Learning Certificate - 23-2475](#)