About this course

Visual representations generated by statistical models help us to make sense of large, complex datasets through interactive exploration, thereby enabling big data to realize its potential for informing decisions. This course covers techniques and algorithms for creating effective visualizations based on principles from graphic design, visual art, perceptual psychology, and cognitive science to enhance the understanding of complex data.

Specific topics covered include:
- Data transformations
- Time series analysis
- Exploratory querying
- Exploratory spatial data analysis
- Statistical graphics

Required prior knowledge and skills

- Basic statistics and computer science knowledge including computer organization and architecture, discrete mathematics, data structures, and algorithms
- Knowledge of high-level programming languages (e.g., C++, Java) and scripting language (e.g., Python)

Learning Outcomes

Learners completing this course will be able to:
- Develop exploratory data analysis and visualization tools using Python and Jupyter notebooks
- Apply design principles for a variety of statistical graphics and visualizations including scatterplots, line charts, histograms, and choropleth maps
- Combine exploratory queries, graphics, and interaction to develop functional tools for exploratory data analysis and visualization

Estimated Workload/Time Commitment Per Week

15 - 20 hours per week
Ross Maciejewski (Dr. Ross) is an Associate Professor at Arizona State University in the School of Computing, Informatics & Decision Systems Engineering and Director of the Center for Accelerating Operational Efficiency, a Department of Homeland Security Center of Excellence. His primary research interests are in the areas of geographical visualization and visual analytics focusing on public health, dietary analysis, social media, criminal incident reports, and the food-energy-water nexus.

Professor Huan Liu joined ASU in 2000 after conducting research in Telecom (Telstra) Australia Research labs and teaching at the National University of Singapore. He has extensive experience in research and development. Liu’s research and teaching focuses on machine learning, data mining, and real world applications.

K. Selcuk Candan is a professor of computer science and engineering at Arizona State University and the director of ASU’s Center for Assured and Scalable Data Engineering (CASCADE). His primary research interest is in the area of management and analysis of non-traditional, heterogeneous, and imprecise (such as multimedia, web, and scientific) data.