Course Overview
Successful business analysts, managers, and executives are increasingly required to make data-driven decisions to run their businesses, rather than rely on experience and intuition alone. This course teaches you the latest data analytic methods and decision methods now used by leading-edge business practitioners, going deep to understand their technical inner workings and going broad to realize their practical business applications.

- Data analysis/business decision methodology
- Data analytic methods, including machine learning and other approaches
- Introduction to R software for data analysis
- Real-world/real-data business practicum across a variety of industries

Suggested Preparation
A background in introductory statistics, introductory business analytics, decision modeling, or programming will be useful, but not required. This course can be taken by itself or in parallel with UGBA 104.

Instructor Biography
Dr. Richard Huntsinger is a professor, Silicon Valley entrepreneur, and Fortune 500 operations executive with broad international experience leading data analytics, enterprise software development, system integration, and operational excellence programs at organizations like Hewlett-Packard, AT&T, Symantec, Hitachi, Exelon, Curtiss-Wright, and US Department of Energy. Recent projects include data analysis and process automation in nuclear power plant operations, electricity demand forecasting in smart electric grid networks, and enterprise data warehouse optimization in high-tech and telecom operations.

Data Analysis Outline

**Data Analysis Methodology**
- Business decision lifecycle, risk profiling, model performance-to-business value analysis

**Data Exploration & Preparation Methods**
- Data selection, 2-D & 3-D data visualization, cross-tabulation, kernel density estimation, imputation, principal component analysis

**Descriptive Data Analytic Methods**
- Hierarchical agglomeration, k-means, Gaussian mixture model by expectation-maximization

**Predictive Data Analytic Methods**
- Naive Bayes, support vector machine, neural network, logistic regression, decision tree, random forest, nearest neighbor, linear regression, ensembles, multinomial classification

**Model Evaluation & Tuning**
- Dissimilarity analysis, confusion matrix analysis, cross-validation, sensitivity analysis, feature tuning, hyper-parameter tuning

**Data Analytic Methods for Special Data Types**
- Time series data, text data

**Data Analysis Software Tools**
- Jupyter, R, ggplot2, rgl

**Labs & Project**
- Stock market, banking, energy, real estate, aviation, hospitality, retail, call center, health care, market segmentation, politics, transportation, telecom, workplace diversity