Data Science (DATA)

DATA 706. Data-Driven Security. 3 Credits.
This course will cover the use of data science techniques such as data preparation, feature selection, exploratory data analysis, visualization and machine learning to efficiently manipulate, analyze and gain valuable insights from cyber security data.

DATA 711. Basic Computational Statistics using R. 3 Credits.
Basic Statistics, General R, Data Manipulation, Basic Statistical Programming Skills, Simple Linear Regression, Classical Testing, and Categorical Data Analysis. Prereq or Coreq: This course is designed for the certificate degree in Big Data Applied Statistics Analysis and it may not be used toward any other statistics degree. Cross-listed with STAT 711.

DATA 712. Applied Statistical Machine Learning. 3 Credits.
This course provides several fundamental concepts and methods in statistical machine learning and big data analysis: divide and conquer, parallel computing in R, linear method for regression, lasso, linear method for classification, logistic regression, KNN, model selection and assessment, regression tree, classification tree, bagging, random forest, boosting, support vector machine, neural networks, K-means clustering, principal components analysis. We use R to implement all the methods in this course. NOTE: It cannot be taken as credit towards M.S. in Applied Statistics or the Ph.D. degree or the Graduate Certificate in Statistics, but may be taken as credit for the Big Data Statistical Analysis Graduate Certificate. This course is also part of the M.S. degree program in Data Science. Cross-listed with STAT 712.

DATA 713. Introduction to Data Science. 3 Credits.
Large Scale Data Manipulation, Data Management, Big Data Construction using Probabilistic and Machine Learning Data Linkage, Web Crawling, Parallel Statistical Computing, and Transferring Data Between Statistical Software. Prereq: STAT 711 or DATA 711. Prereq or Coreq: This course is designed for the certificate degree in Big Data Applied Statistics Analysis and it may not be used toward any other statistics degree. Cross-listed with STAT 713.

DATA 714. Statistical Big Data Visualization. 3 Credits.
This course is designed to equip students with the theoretical and practical tools needed to build effective and engaging data visualizations and demonstrate competence in designing and developing visual stories with data. Students will learn visual representation methods and techniques that improve their understanding of complex data and models, with emphasis placed on the identification of patterns, trends, and differences from data sets across categories, space, and time using R, PowerBI, and Tableau. NOTE: It cannot be taken as credit towards M.S. in Applied Statistics or the Ph.D. degree or the Graduate Certificate in Statistics, but may be taken as credit for the Big Data Statistical Analysis Graduate Certificate. This course is also part of the M.S. degree program in Data Science. Prereq: STAT 711. Dual-listing: STAT 714.

DATA 720. Programming for Data Science. 3 Credits.
This course introduces the foundations of programming for data science. It covers basics of programming in python and python tools for data acquisition, wrangling, analysis, and visualization. This course is for non-CS major students.

DATA 725. Applied Statistics. 3 Credits.
Data description, probability, inference on means, proportions, difference of means and proportions, categorical data, regression, analysis of variance, and multiple comparisons. This course is not intended for statistics or mathematics majors. Cross-listed with STAT 725.

DATA 726. Applied Regression and Analysis of Variance. 3 Credits.
Simple and multiple regression, ANOVA tables, correlation, regression diagnostics, selection procedures, analysis of covariance, one-way ANOVA, two-way ANOVA. Prereq: STAT 725 or DATA 725. Cross-listed with STAT 726.

DATA 760. Applied Artificial Intelligence. 3 Credits.
Since Artificial intelligence (AI) is a very broad and diverse field that mimics human intelligence using computer programs and algorithms, an introduction to the field and the applied AI perspective will be discussed. Different AI techniques applied to different real-world problems will be the focus of the course.

DATA 761. Applied Machine Learning. 3 Credits.
This course is for Data Science students and will introduce applied machine learning methods (supervised and unsupervised learning) through the use of python packages. Topics will include classification algorithms and unsupervised clustering techniques. Prereq: DATA 720.

DATA 765. Applied Database Systems. 3 Credits.
Introduction to database systems focusing on relational database. The topics include Introduction to various database models, Data modeling, Database theory, Database design principles, SQL (structured query language) statement, Develop applications with database, and Database management. Prereq: DATA 720.