MASTER OF SCIENCE IN DATA SCIENCE AND STATISTICS

Program Director

Dr. G. Jay Kerns
620 Lincoln Building
(330) 941-3310
gkerns@ysu.edu

Program Description

The Department of Mathematics and Statistics offers the M.S. degree in Data Science and Statistics traditionally. Concentrations in this degree include:

- data science,
- statistics,
- GIS, and
- data analytics.

In this collaborative program, graduate faculty members have a broad range of research interests in data science, statistics, and applications to domain fields. The curriculum stresses theoretical as well as computational aspects and is flexible enough to key a student's program to individual interests and abilities.

Admission Requirements

The admission requirements are those specified as the minimum admission requirements of the College of Graduate Studies, which can be found at: https://catalog.ysu.edu/graduate/admission/ . Students not satisfying all of these requirements may be admitted with provisional status subject to the approval of the graduate program director and the graduate dean.

Jozsi Z. Jalics, Ph.D., Professor
Computational neuroscience; mathematical biology; dynamical systems; partial differential equations

G. Jay Kerns, Ph.D., Professor
Signed measures; infinite divisibility; exchangeability in probability and statistics; applications of stochastic processes

Lucy Xiaojing Kerns, Ph.D., Associate Professor
Simultaneous confidence bands; minimum effective doses; benchmark dose methodology

Thomas L. Madsen, Ph.D., Associate Professor
Abstract algebra; group theory; representation theory

Nguyet Thi Nguyen, Ph.D., Associate Professor
Financial models; Monte Carlo simulation; actuarial science

Anita C. O'Mellan, Ph.D., Professor
Graph theory; combinatorics; early childhood mathematics education

Alicia Prieto Langarica, Ph.D., Professor
Mathematical biology; agent-based modeling

Thomas Smotzer, Ph.D., Professor
Real analysis; measure theory; operator theory

Jamal K. Tartir, Ph.D., Professor
Set-theoretic topology

Padraic ("Paddy") W. Taylor, Ph.D., Associate Professor
Multipoint Boundary Value Problems

Thomas P. Wakefield, Ph.D., Professor, Chair

Character theory; actuarial science

- A minimum of 30 semester hours of credit
- A cumulative grade point average of at least 3.0
- The student must complete core degree requirements comprising the following courses or their equivalent:

COURSE	TITLE	S.H.
DATX 5801	Data Management	3
DATX 6903	Data Visualization	3
DATX 6905	Predictive Modeling Algorithms	3
PHIL 6926	Data Ethics	3
STAT 6940	Advanced Data Analysis	3
Choose one of the following:		3-8
One year commitment to the YSU Data Mine (DATX 5895 and 6996)		
or		
DATX 6996	Data Analytics Project	
or STEM 6998	STEM Graduate Internships	
Electives (see list below)		12
Total Semester Hours		**30-35**

- Students are strongly encouraged to participate for one-year in the YSU Data Mine as their culminating experience.
- At least 15 hours of the student's approved program must be at the 6900 level. In addition to completing the courses which make up the core, students must complete additional hours of elective courses to satisfy 30-semester hour requirement for the degree. Recommended course groupings are described below.
- Before completing 12 semester hours, the student must submit the entire degree program for approval and evaluation by the Graduate Executive Committee. Subsequent revisions to this program must be approved by the Graduate Executive Committee.
- Students must participate in an exit interview during the semester in which they plan on graduating. The exit interview will be conducted with one or more members of the Graduate Executive Committee.

Electives

Students satisfy the elective requirement for the degree by choosing a courses from the following list. Other courses may be selected subject to approval of the Graduate Executive Committee.

COURSE	TITLE	S.H.
BIOL 5858	Computational Bioinformatics	3
BIOL 6900	Advanced Bioinformatics	3
CSCI 6950	Advanced Database Design and Administration	3
CSCI 6951	Data Science and Machine Learning	3
CSCI 6952	Deep Learning	3
CSCI 6970	Biometrics	3
CSCI 6971	Cloud Computing and Big Data	3
DATX 5800	Quantitative Methods in Economic Analysis	3
ECON 6976	Econometrics	3
GEOG 6901	Introduction to Geographic Information Science	3
GEOG 6902	Introduction to Remote Sensing	3
GEOG 6903	Advanced Geographic Information Science	3
GEOG 6904	Advanced Remote Sensing	3
ISEN 6902	Digital Simulation	3
ISEN 6935	Decision Analysis for Engineering	3
MPH 6904	Biostatistics in Public Health	3
MATH 5835	Introduction to Combinatorics and Graph Theory	3
MATH 5845	Operations Research	3

MATH 6910	Advanced Engineering Mathematics 1	3
MATH 6911	Advanced Engineering Mathematics 2	3
STAT 5811	SAS Programming for Data Analytics	3
STAT 5814	Statistical Data Mining	3
STAT 5819	Bayesian Statistics	3
STAT 5840	Statistical Computing	3
STAT 5846	Categorical Data Analysis	3
STAT 5849	Multivariate Statistical Analysis	3
STAT 5857	Statistical Consulting	3
STAT 5895	Special Topics in Statistics	2-3
STAT 6904	Actuarial Mathematics 1	3
STAT 6905	Actuarial Mathematics 2	3
STAT 6910	Advanced Short-Term Actuarial Mathematics	3
STAT 6911	Advanced Long-Term Actuarial Mathematics	3
STAT 6912	Advanced SAS Programming for Data Analytics	3
STAT 6943	Mathematical Statistics 1	3
STAT 6944	Mathematical Statistics 2	3
STAT 6948	Linear Models	3
STAT 6949	Design and Analysis of Experiments	3
ECON 6915	Health Care Analytics	3

Students with particular interests or career goals are advised to choose their elective courses based upon the recommendations below.

Data Science

COURSE	TITLE	S.H.
CSCI 6950	Advanced Database Design and Administration	3
CSCI 6951	Data Science and Machine Learning	3
CSCI 6952	Deep Learning	3
CSCI 6971	Cloud Computing and Big Data	3
MATH 5835	Introduction to Combinatorics and Graph Theory	3

Statistics

COURSE	TITLE	S.H.
STAT 5811	SAS Programming for Data Analytics	3
STAT 5814	Statistical Data Mining	3
STAT 5819	Bayesian Statistics	3
STAT 5840	Statistical Computing	3
STAT 5846	Categorical Data Analysis	3
STAT 5849	Multivariate Statistical Analysis	3
STAT 5857	Statistical Consulting	3
STAT 5895	Special Topics in Statistics	2-3
STAT 6912	Advanced SAS Programming for Data Analytics	3
STAT 6943	Mathematical Statistics 1	3
STAT 6944	Mathematical Statistics 2	3
STAT 6948	Linear Models	3
STAT 6949	Design and Analysis of Experiments	3

GIS

COURSE	TITLE	S.H.
GEOG 6901	Introduction to Geographic Information Science	3
GEOG 6902	Introduction to Remote Sensing	3
GEOG 6903	Advanced Geographic Information Science	3
GEOG 6904	Advanced Remote Sensing	3

BioInformatics

COURSE	TITLE	S.H.
CSCI 6970	Biometrics	3
BIOL 5858	Computational Bioinformatics	3
BIOL 6900	Advanced Bioinformatics	3
MATH 6910	Advanced Engineering Mathematics 1	3
MATH 6911	Advanced Engineering Mathematics 2	3

Business Analytics

COURSE	TITLE	S.H.
DATX 5800	Quantitative Methods in Economic Analysis	3
ECON 6976	Econometrics	3
ISEN 6902	Digital Simulation	3
ISEN 6935	Decision Analysis for Engineering	3
MATH 5845	Operations Research	3

Accelerated MS Data Science

Undergraduate students can apply for admission into the accelerated program for the MS in Data Science and Statistics after completing 78 semester hours with a GPA of 3.3 or higher. After being admitted into the program, students can take a maximum of nine semester hours of graduate coursework that can count toward both a bachelor's and master's degree. The courses chosen to count for both undergraduate and graduate coursework must be approved by the Graduate Executive Committee upon admission into the program. An additional three hours of graduate coursework can be completed as an undergraduate and used exclusively for graduate credit.

Learning Outcomes

Students will manipulate and prepare large data sets for analysis through common techniques to clean data and identify trends and outliers.

Students will develop an ethical framework from which to critically examine the origins, uses, and implications of their work with data.

Students will learn to describe and apply the common techniques used in statistics and predictive modeling and choose an appropriate technique to model and to make predictions on a dataset.

Students will demonstrate that they can communicate data-driven results effectively, both orally and in writing, by completing a graduate project, internship or through participation in the YSU Data Mine.

DATX 5800   Quantitative Methods in Economic Analysis   3 s.h.
This course introduces to students the nuts and bolts of cleaning, manipulating, and crunching data in Python, and serves as adequate preparation to enable students to move on to other domain-specific courses that use Python as the learning tool.
Prereq.: STAT 2601 or STAT 2625 or STAT 3717 or STAT 3743 or ECON 3790, or ECON 3788 and ECON 3789, or ECON 3788 and BUS 3700, or permission of instructor.

DATX 5801   Data Management   3 s.h.
This course covers the basic concepts of database systems and emphasizes the real-world database applications relevant to the management of data in an organization environment. The topics include (not limited to) database environment, database development, relational database management systems, SQL/NoSQL data management language, data normalization, data warehousing, and internet database environment. Credit will not be given for both DATX 5801 and CSIS 3722.
Prereq.: Junior standing or higher and GPA of 2.5 or higher.

DATX 5803   Data Visualization   3 s.h.
Data visualization refers to the graphical representation of information revealed through data analysis. With the assistance of various visualization elements, we can present data in a clear and effective manner. More importantly, turning data into impactful images, we are able to gain valuable insights and intelligence that help improve our decision-making processes. This course introduces students to various types of visualization techniques like charts, tables, graphs, maps, infographics and dashboards. It emphasizes applying appropriate visualization techniques in uncovering information from data. Moreover, it will help students develop skills of data storytelling, i.e. effectively communicating actionable insights through the combination of data visualization and narratives.
Prereq.: Junior standing or higher and GPA of 2.5 or higher.

DATX 5805   Predictive Modeling Algorithms   3 s.h.
Predictive modeling (also referred to predictive analytics and machine learning) applies statistical techniques in analyzing data to predict outcomes. Through a hands-on approach, this course helps students develop basic skills in predictive analytics. Topics may include (not limited to) k-nearest neighbors, naïve-Bayes, linear and logistic regression models, time-series models, classification and regression trees, Principle Component/Factor Analysis, non-linear models, neural networks, random forests, and cluster analysis among others.
Prereq.: Junior standing or higher and GPA of 2.5 or higher.

DATX 5895   Selected Topics in Data Analytics   1-3 s.h.
The study of a topic in data analytics in depth or the development of a special area of data analytics. May be repeated with permission of the instructor.
Prereq.: Permission of the instructor.

DATX 5896   Data Analytics Project   3 s.h.
Individual research project culminating in a written report or paper utilizing predictive modeling techniques, visualization, and data management techniques. May be repeated with permission of instructor.
Prereq.: Permission of instructor.
Coreq.: DATX 5895.

DATX 5896C   CE Data Analytics Project   3 s.h.
Individual research project culminating in a written report or paper utilizing predictive modeling techniques, visualization, and data management techniques. May be repeated with permission of instructor.
Prereq.: Permission of instructor.
Coreq.: DATX 5895.

DATX 6903   Data Visualization   3 s.h.
This course introduces students to various types of visualization techniques such as charts, tables, graphs, maps, infographics and dashboards. It emphasizes applying appropriate visualization techniques in uncovering information from data. Moreover, it will help students effectively communicate actionable insights through the combination of data visualization and narratives. Credit will not be given for both DATX 5803 and DATX 6903.
Prereq.: Graduate Standing.

DATX 6905   Predictive Modeling Algorithms   3 s.h.
Predictive modeling (also referred to predictive analytics and machine learning) applies statistical techniques in analyzing data to predict outcomes. Through a hands-on approach, this course helps students develop basic skills in predictive analytics. Topics may include (not limited to) k-nearest neighbors, naïve-Bayes, linear and logistic regression models, time-series models, classification and regression trees, Principal Component/Factor Analysis, non-linear models, neural networks, random forests, and cluster analysis among others. Credit will not be given for both DATX 5805 and DATX 6905.
Prereq.: Graduate Standing.

DATX 6995   Selected Topics in Data Analytics   1-3 s.h.
The study of a topic in data analytics in depth or the development of a special area of data analytics. May be repeated with permission of the instructor.
Prereq.: Permission of the instructor.

DATX 6996   Data Analytics Project   1-3 s.h.
Individual or team research project culminating in a written report or paper utilizing predictive modeling techniques, visualization, and data management techniques, possibly through a partnership with a business, industry, or government partner. If working in partnership with YSU Data Mine, concurrent enrollment in DATX 5895 is required. May be repeated.
Prereq.: Permission of instructor.