Applied Statistics and Data Science (MS)

Graduates of the **Master of Science (MS) in Applied Statistics and Data Science** will be trained in the data science process, machine learning, data visualization, statistical inference, algorithmic and computational thinking, experimental design, coding, ethics, and algorithmic accountability. Moreover, they will acquire competency in the following areas.

- Computational and statistical thinking.
- Mathematical foundations.
- Algorithms and software foundation.
- Data curation.
- Knowledge transference—communication and responsibility.

Why UTRGV?

- Ranked #79 among 300+ national universities by Washington Monthly in 2018
- Ranked 4th best college in Texas 2018 by BestColleges.com
- Accredited, cutting edge degree program
- Experienced, dedicated faculty
- Affordable tuition (ranked #1 in net price among national universities by Washington Monthly in 2018 and #3 most affordable university in America 2018 by BestValueSchools.com)
- Demonstrated student success in research, professional certification and career advancement

Admission Requirements

Step #1: Submit a UTRGV Graduate Application at www.utrgv.edu/gradapply. There is no application fee.

Step #2: Request your official transcripts to be sent electronically to gradapps@utrgv.edu or mailed to:

The University of Texas Rio Grande Valley
The Graduate College
Marialice Shary Shivers Bldg. 1.158
1201 W. University Drive
Edinburg, TX 78539-2999

***Please Note:** If you are a graduate of UTPA, UTB/TSC, or UTRGV you do not need to request an official transcript to be sent to the Graduate College.

Review and submit all Program Requirements:

- Bachelor's degree in any field with a minimum of 12 hours of upper-division mathematics or statistics course work.
- Undergraduate GPA of at least 3.0 in upper-level Mathematics and/or Statistics courses.
- Official transcripts from each institution attended (must be submitted directly to UTRGV).
- Letter of Intent detailing professional goals and reasons for pursuing the graduate degree.

Additional requirements for domestic applicants who attended foreign universities:

- TOEFL or IELTS Language Proficiency Test with minimum scores: 550 on paper-based, 213 on computer based, or 79 on Internet-based for the TOEFL; 6.5 for the IELTS. TOEFL and IELTS scores are valid for 2 years. For additional information, visit the Additional Documents for Domestic Applicants who Attend Foreign Universities section of our website.
- Certified English translation of educational records.

Additional requirements for international applicants:

- TOEFL or IELTS Language Proficiency Test with minimum scores: 550 on paper-based, 213 on computer based, or 79 on Internet-based for the TOEFL; 6.5 for the IELTS. For additional information, visit the *English Proficiency Exam* section of our website.
- Certified English translation of educational records.
- Financial Documentation showing sufficient funds to cover all expenses (living and academic) for the first year of study. For additional information, visit the *Financial Documentation* section of our website.
- Immigration documents, including a current copy of your valid passport. For additional information, visit the *Immigration Documents* section of our website.

UPDATE ON INTERNATIONAL ADMISSIONS FROM U.S. IMMIGRATION AND CUSTOMS ENFORCEMENT:

- SEVP regulations prohibit the issuance of a Form I-20 based on **conditional admission,** effective July 13, 2016. University officials can only issue a Form I-20 when students have met all standards for admission for the program of study listed on the Form I-20. These standards for admission include any English proficiency requirements.

Program Contact

Program Coordinator: Dr. George Yanev

Phone: (956) 665-3632

Office: Edinburg Campus, EMAGC 2.316

E-Mail: george.yanev@utrgv.edu

Deadlines

Deadlines:

Applications will be accepted year round and prospective students are encouraged to apply at least 2 months before classes start to ensure a timely application review. Applying early will also give prospective students the best opportunity to be considered for scholarships and other possible funding opportunities.

***Note**: This program is schedule to start in **Fall 2020** and will admits applicants during **Fall, Spring, Summer I and Summer II** semesters.

Tuition Estimate

Residency	Per 3-Credit Hour Course	Semester (9-Credit Hours)	Total Estimated Cost
Texas Resident	$1,481.39	$3,894.17	$17,776.68
Non-Resident/International	$2,708.39	$7,575.17	$32,500.68

**We estimate that tuition and fees will closely approximate the rates shown above; however, rates are subject to change. Please note that the rates above are estimated for on-campus students and those enrolled in 16-week online programs. The rate is different for Accelerated Online Programs (AOP). Visit the tuition and fees page for detailed information.*

Course Requirements

Required Courses	**21**
MATH 6330: Linear Algebra	3
MATH 6364: Statistical Methods	3
MATH 6365: Probability and Statistics	3
MATH 6392*: Statistical Learning	3
CSCI 6302: Foundations of Software and Programming Systems for Information Technology	3
CSCI 6305: Foundations of Algorithms and Programming Languages	3
CSCI 6366: Data Mining and Warehousing	3

Prescribed Electives	**9**
Statistics Courses (choose one)	
MATH 6336: Advanced Sampling	3
MATH 6379: Stochastic Processes	3
MATH 6380: Time Series Analysis	3
MATH 6381: Mathematical Statistics	3
MATH 6382: Statistical Computing	3
MATH 6383: Experimental Design and Categorical Data	3
MATH 6384: Biostatistics	3
Computer Science Courses (choose one)	
CSCI 6315: Applied Database Systems	3
CSCI 6333: Advanced Database Design and Implementation	3
CSCI 6352*: Advanced Machine Learning	3
CSCI 6350: Advance Artificial Intelligence	3
CSCI 6355: Bioinformatics	3

Mathematics Courses (choose one)

MATH 6352: Analysis I	3
MATH 6375: Numerical Analysis	3
Capstone Requirement	**6**
Thesis Option	
MATH 7300: Thesis I	3
MATH 7301: Thesis II	3
Master Project Option	
MATH 6390* or MATH 6391: Internship* or Master Project	3
MATH XXXX or CSCI XXXX: Prescribed Elective	3
Non-Thesis Option (Comprehensive Exam)	
MATH XXXX or CSCI XXXX: Prescribed Elective	3
MATH XXXX or CSCI XXXX: Prescribed Elective	3
Total hours required for completion:	**36**

Course Descriptions:

CSCI 6302: Foundations of Software and Programming Systems for Information Technology [3-0]

Focusing on a high level object oriented language (e.g., Java, C++), provides foundational study of algorithms, data structures, and programming systems in the context of information technology systems. **Prerequisite**: Knowledge of a high level programming language and consent of instructor.

CSCI 6305: Foundations of Algorithms and Programming Languages [3-0]

In-depth analysis of computing algorithms and data structures for implementation in the context of software engineering design using structured programming languages. **Prerequisites:** CSCI 6302 or CSCI 6382.

CSCI 6315: Applied Database Systems [3-0]

Course covers the application of a modern database system. Concepts covered include relational model, normalization, structured query language, Internet data formats, and server and client side technologies. The course is targeted at students who are interested in the development of application programs usinga database system such as Oracle, Teradata, or Microsoft SQL. **Prerequisites**: CSCI 6302 or equivalent.

CSCI 6333: Advanced Database Design and Implementation [3-0]

Focuses on distributed database systems. Includes file allocation, directory systems, deadlock detection a prevention, synchronization, query optimization, and fault tolerance. The course will include one or more programming projects demonstrating implementation of concepts introduced. **Prerequisite**: CSCI 6305, or

consent of instructor.

CSCI 6350: Advance Artificial Intelligence [3-0]

Issues of knowledge representation, including a survey of important knowledge-based systems. Current research issues, including neural networks, object-oriented programming in AI, natural language understanding, device understanding, and perception. **Prerequisite**: CSCI 6305 or consent of instructor.

CSCI 6352: Advanced Machine Learning [3-0]

This course provides selected topics in Machine Learning focusing on Deep Learning. Topics include: logistic regression, neural networks, deep neural networks, convolutional neural networks, recurrent neural networks, reinforcement learning, regularization, Bayesian network classifiers, and clustering techniques. The course also explores the applications of such topics in Data Science, Data Mining, and Data Analytics. Through projects, the taught algorithms are used to build intelligent systems in areas such as smart robots, computer vision, medical and healthcare informatics.

CSCI 6355: Bioinformatics [3-0]
Examines the creation and development of advanced information and computational techniques for problems in the biosciences, including biology, biochemistry, biotechnology, and medicine. Presents advanced concepts and techniques of bioinformatics and computational biology tools to solve problems in topics such as sequence alignment, gene and motif finding, restriction mapping, microarray data analysis and gene expressions.
Prerequisites: CSCI 6305 or consent of instructor

CSCI 6366: Data Mining and Warehousing [3-0]

As a multidisciplinary field, draws on work from areas including database technology, artificial intelligence, machine learning, neural network, statistics, information retrieval, and data visualization. Theoretical and practical methods will be presented on knowledge discovery and systems design and implementation. **Prerequisite:** CSCI 6305.

MATH 6330: Linear Algebra [3-0]

Topics include the proof-based theory of matrices, determinants, vector spaces, linear spaces, linear transformations and their matrix representations, linear systems, linear operators, eigenvalues and eigenvectors, invariant subspaces of operators, spectral decompositions, functions of operators and applications to science, industry and business. **Prerequisite**: MATH 2318 Linear Algebra with a grade of "C" or higher.

MATH 6336: Advanced Sampling [3-0]

This course will focus on planning, execution and analysis of sampling from finite populations; simple, stratified, multistate and systematic sampling; ratio estimates. Departmental approval.

MATH 6352: Analysis I [3-0]

The purpose of this course is to provide the necessary background for all branches of modern mathematics involving analysis and to train the student in the use of axiomatic methods. Topics include metric spaces, sequences, limits, continuity, function spaces, series, differentiation and the Riemann integral. **Prerequisite**: MATH 3372 Real Analysis I with a grade of "C" or higher.

MATH 6364: Statistical Methods [3-0]

This is a course in the concepts, methods and usage of statistical data analysis. Topics include test of hypotheses and confidence intervals; linear and multiple regression analysis; concepts of experimental design, randomized blocks and factorial analysis; a brief introduction to non-parametric methods; and the use of statistical software. Departmental approval.

MATH 6365: Probability and Statistics [3-0]

Topics in this course include set theory and concept of probability, conditional probability, random variables,

discrete and continuous probability distributions, distribution and expectations of random variables, moment generating functions, transformation of random variables, order statistics, central limit theorem and limiting distributions. **Prerequisite**: MATH 2415 Calculus III with a grade of "C" or higher, or consent of instructor.

MATH 6375: Numerical Analysis [3-0]

This course provides a fundamental introduction to numerical techniques used in mathematics, computer science, physical sciences and engineering. The course covers basic theory on classical fundamental topics in numerical analysis such as: computer arithmetic, approximation theory, numerical differentiation and integrations, solution of linear and nonlinear algebraic systems, numerical solution of ordinary differential equations and error analysis of the abovementioned topics. Connections are made to contemporary research in mathematics and its applications to the real world. **Prerequisite**: MATH 2318 Linear Algebra and 2415 Calculus III with grades of "C" or better; and computer programming; or consent of instructor.

MATH 6379: Stochastic Processes [3-0]

Discrete and Continuous Time Markov Processes, Poisson Processes, Renewal Processes, Diffusion Processes, Brownian Motion. **Prerequisite**: MATH 6365.

MATH 6380: Time Series Analysis [3-0]

This course is an introduction to statistical time series analysis. Topics include: ARIMA and other time series models, forecasting, spectral analysis, and time domain regression. Model identification, estimation of parameters, and diagnostic checking are included. **Prerequisite**: MATH 6379.

MATH 6381: Mathematical Statistics [3-0]

This course in mathematical Statistics includes theory of estimation and hypothesis testing; point estimation, interval estimation, sufficient statistics, decision theory, most powerful tests, likelihood ration tests, chi-square tests, minimum variance estimation, Neyman-Pearson theory of testing hypotheses, elements of decision theory. **Prerequisite**: MATH 6365.

MATH 6382: Statistical Computing [3-0]

A course in modern computationally-intensive statistical methods including simulation, optimization methods, Monte Carlo integration, maximum likelihood /EM parameter estimation, Markov chain Monte Carlo methods, resampling methods, non-parametric density estimation. Consent of instructor.

MATH 6383: Experimental Design and Categorical Data [3-0]

Design and analysis of experiments, including one-way and two-way layouts; factorial experiments; balanced incomplete block designs; crossed and nested classifications; fixed, random, and mixed models; split plot designs, inference for categorical data, contingency tables, generalized linear models, logistic regression, logit and loglinear models. **Prerequisite**: MATH 6364.

MATH 6384: Biostatistics [3-0]

This course is a survey of crucial topics in biostatistics; application of regression in biostatistics; analysis of correlated data; logistic and Poisson regression for binary or count data; survival analysis for censored outcomes; design and analysis of clinical trials; sample size calculation by simulation; bootstrap techniques for assessing statistical significance; data analysis using R. Consent of instructor.

MATH 6390: Internship [3-0]

In this intership students will seek to apply the knowledge they acquired in the program to an intership in the private or government sector. They will have the opportunity to gain insight and experience in applying statistics and data science principles and concepts in an actual work-related environment. The students will perform the

internship under the supervision of both a statistics faculty member and a collaborating member of the

participating intership site. **Prerequisite**: Consent of graduate program director.

MATH 6391: Master Project [3-0]

Individual work or research on advanced mathematical problems conducted under the direct supervision of a faculty member.

MATH 6392: Statistical Learning [3-0]

This course introduces the statistical tools of supervised and unsupervised learning including topics of regression and clssification, such as linear regression, multiple regression, logistic regression, K-nearest neighbors, polynomial regression, splines regression, tree regression, random forests, ridge regression and the Lasso, linear and quadratic discriminant analysis, support vector classifiers and machines. In this course, students are required to use a statitiscal software to carry out the procedures. **Prerequisite**: Consent of instructor.

MATH 7300: Thesis I [3-0]

First part of two course sequence. Graduate standing and consent of thesis advisor.

MATH 7301: Thesis II [3-0]

Second part of two course sequence. Graduate standing and consent of thesis advisor.