Department of Statistics and Data Science

Program of Data Science and Big Data Technology for International Students (2024)

I. Introduction

Established in April 2019, the Department of Statistics and Data Science aims to build up a world-class educational training and research center. The department is committed to cultivating top-notch talents with solid scientific knowledge, active thoughts, innovative awareness, and global vision. Until June 2023, the department has 18 full-time and 4 jointly appointed faculty members, including 3 Chair Professors, 4 Professors, 5 Associate Professors, 10 Tenure-track Assistant Professors. All faculty members have extensive overseas study or work experiences. One member is an invited speaker at the International Congress of Mathematics, and IMS Medallion Lecturer. Two members are the winners of the prestigious State Natural Science Award (2nd class). At present, the department has two undergraduate programs, namely the Program of Statistics and the Program of Data Science and Big Data Technology, as well as two graduate programs (M.Phil. and Ph.D.), which cover a broad array of research areas including Biostatistics, Clinical Trial Design, High Dimensional Data Analysis, Random Matrix, Time Series Analysis, Bayesian Statistics, Financial Statistics, Limit Theory in Probability and Statistics, Data Science and Big Data Technology. Statistics is applied extensively in various disciplines, from natural sciences (like physics, chemistry, biomedicine, etc.) to social sciences and humanities, as well as business and government decision-making. The undergraduate program of Statistics focuses on applying probability theory to establish statistical models based on the collected data, conduct quantitative analysis, and make inferences and predictions to serve as the reference for decision-making.

Academic subject areas: Data Science and Big Data Technology

Professional code: 080910T

II. Objectives and Learning Outcomes

1. Objectives

The objective for international undergraduates majoring in Data Science and Big Data Technology is to cultivate professional talents who are interested in data analysis and big data technology. International undergraduates in this major will have a solid theoretical foundation in mathematics, theoretical foundations in statistics and big data technology, skilled computer programming techniques, good at practical data collection, mining, modeling and analysis, algorithm design and data visualization, and

strong social communication skills.

Moreover, they will be able to conduct further research related to artificial intelligence or engage in data analysis, data mining, information management in enterprises and government departments. Graduates can pursue further graduate studies as well as employment in a wide range of industries and sectors: government departments that need to engage in data analysis; financial departments such as banks, securities, and fund companies; communication, software, and robotics industries; market analysis, risk management, and quality management departments of small, medium, and large enterprises; higher and secondary schools and research institutes; survey companies, consulting companies, etc. Specialized agencies for data analysis, and various industrial sectors that generate demand for big data processing and analysis, etc. In the era of big data, data science faces a wealth of opportunities and challenges. Graduates of Data Science and Big Data Technology major will have a strong theoretical background and a broad range of knowledge to seize the opportunities and meet the challenges.

2. Learning Outcomes

2.1 Students should have a solid mathematical foundation, master the basic knowledge and theories of statistics, and understand the basics of computer science, etc.

2.2 Students should be able to proficiently read data science literature in English, master the principal methods of literature search, information retrieval, and data query with modern information technology.

2.3 Students should be equipped with the essential skills of applying data science knowledge and principles to analyze and solve practical problems.

III. Study Length, Degree and Graduation Requirements

1. Study Length: Four years.

2. Degree Conferred: Students who complete and meet the degree requirements of the undergraduate program will be awarded a bachelor's degree in science.

3. The Minimum Credit Requirement for Graduation: 151 credits. The specific requirements are as follows.

Module		**Category**	**Minimum Credit Requirement**
General Education Courses	Chinese Language and Culture Module	Chinese Language and Culture	16
	Arts and Physical Education Module	Physical Education	4
		Arts	2
	Competence Development Module	Computer Programming	3

		Writing	2
		Foreign Languages	14
	Humanities and Social Sciences Module	Humanities	6
		Social Sciences	
		Chinese Studies	2
	Mathematics and Natural Sciences Module	Mathematics	12/14
		Physics	10
		Chemistry	3
		Geoscience + Life Science	3
	GE to Majors Bridging Module	Introduction to Majors	2
Major Courses	Major Required Courses	Major Foundational Courses	20
		Major Core Courses	18
		Practice-based Learning (Undergraduate Thesis, Internships, Research Projects, etc.)	14
	Major Elective Courses	Major Elective Courses	20/18
Total Credits			151
Note: Please see the General Education Requirement for more details on Chinese Language and Culture Module, Arts and Physical Education Module, Competence Development Module（Foreign Languages & Writing）, Humanities and Social Sciences Module, and GE to Majors Bridging Module.			

IV. Course Requirements for the Mathematics and Natural Sciences Module and Computer Programming

Course Category	Course Code	Course Name	Credit	Terms	Prerequisite	Department
Mathematics	MA101a / MA117	Mathematical Analysis I / Calculus I	5/4	1 Fall	None	Mathematics
	MA102a / MA127	Mathematical Analysis II / Calculus II	5/4	1 Spring	MA101a/ MA117	Mathematics
	MA107 / MA113	Advanced Linear Algebra I / Linear Algebra	4	1 Fall / 1 Fall-Spring	None	Mathematics
Physics	PHY101/ PHY105	General Physics I / College Physics I	5/4	1 Fall	None	Physics
	PHY102/ PHY106	General Physics II/ College Physics II	5/4	1 Spring	PHY101/ PHY105	Physics
	PHY104B	Experiments of Fundamental Physics	2	1-2 Spring & Fall	None	Physics
Chemistry	CH103/ CH105	General Chemistry/ Chemistry: The Central Science	3	1-2 Spring & Fall	None	Chemistry
Geoscience + Life Science	BIO103/ BIO102B/ EOE100	Principles of Biology/ Introduction to Life Science/ Introduction to Earth Sciences	3	1-2 Spring & Fall	None	Biology/ ESS, OCE, ESE
Computer Programming	CS109/ CS110/	Introduction to Computer Programming/ Introduction to Java Programming/	3	1-2 Spring & Fall	None	Computer Science and Engineering

V. Prerequisites for Major Declaration

Major Declaration Time	Course Code	Course Name	Prerequisite
Declare Major at the End of the First Academic Year	MA101a/MA117	Mathematical Analysis I	
	MA102a/MA127	Mathematical Analysis II	MA101a/MA117
	MA107 / MA113	Advanced Linear Algebra I / Linear Algebra	
	PHY101 / PHY105	General Physics I / College Physics I	
	PHY102 / PHY106	General Physics II / College Physics II	PHY101 / PHY105
	CS109/ CS110/ CS111/ CS112	Introduction to Computer Programming/ Introduction to Java Programming/ Introduction to C Programming/ Introduction to Python Programming	
Declare Major at the End of the Second Academic Year	MA101a / MA117	Mathematical Analysis I / Calculus I	None
	MA102a / MA127	Mathematical Analysis II / Calculus II	MA101a / MA117
	MA107 / MA113	Advanced Linear Algebra I / Linear Algebra	None
	PHY101 / PHY105	General Physics I / College Physics I	None
	PHY102 / PHY106	General Physics II / College Physics II	PHY101 / PHY105
	PHY104B	Experiments of Fundamental Physics	None
	CH103 / CH105	General Chemistry / Chemistry: The Central Science	None
	BIO103/ BIO102B/ EOE100	Principles of Biology/ Introduction to Life Science/ Introduction to Earth Sciences	None
	CS109/ CS110/	Introduction to Computer Programming/ Introduction to Java Programming/	None

Note:

1. If the number of students entering a major at the end of the first academic year in the department is greater than or equal to the total number of the teaching-research faculty (PI)*2*60%, all majors in the department may implement the prerequisites for major declaration at the end of the second academic year.
2. If the number of students entering a major at the end of the first academic year in the department is less than the total number of the teaching-research faculty (PI)*2*60%, all majors in the department do not implement the prerequisites for major declaration at the end of the second academic year.
3. Suppose the number of students applying for a major at the end of the first academic year exceeds four times the total number of the teaching-research faculty (PI), then the department may select students according to predetermined rules. In principle, the rules set by the department shall examine the students' suitability for the major and not based on weighted GPA (Specific rules shall be set by the department and announced in advance).
4. For departments that do not implement prerequisites for major declaration at end of the second academic year, if the cumulative number of students applying for a major at the end of the second academic year and the number of students who have entered a major at the end of the first academic year exceeds four times the total number of the teaching-research faculty (PI), the department may select students according to predetermined rules. In principle, the rules set by the department shall examine the students' suitability for the major and not based on weighted GPA (Specific rules shall be set by the department and announced in advance).

VI. Major Course Arrangement

Table 1: Major Required Courses

Course Category	Course Code	Course Name	Credits	Practice-based Learning Credits	Terms	Prerequisite	Dept.
Major Foundational Courses	STA203	Fundamentals of probability theory	3		2/ Autumn	MA102a/MA127 and MA113	Statistics
	MA204	Mathematical Statistics	3		2/ Spring 3/ Autumn	MA215/MA212	Statistics
	STA201	Operational Research and Optimization	3		2/ Spring	MA107/MA113	Statistics
	CS203/CS203B	Data Structures and Algorithm Analysis / Data Structures and Algorithm Analysis B	3	1	2/ Spring	CS109	Computer Science
	STA204	Discrete Mathematics	3		2/Autumn	MA127 and MA113	Computer Science / Mathematics
	MA203a /MA213-16	Mathematical Analysis III/ Mathematical analysis	5		2/ Autumn	MA102a/MA127	Mathematics
	Total		20				
Major Core Courses	STA321	Distributed Storage and Parallel Computing	3	1	3/ Autumn	CS109/CS110, CS203/CS203B	Statistics
	STA303	Artificial Intelligence B	3	1	3/ Autumn	CS109/CS110, CS203/CS203B;	Computer Science
	MA329	Statistical linear model	3		3/ Autumn	MA204/MA212	Statistics
	MA304	Multivariate Statistical Analysis	3		3/ Spring	MA204/MA212	Statistics
	STA323	Big Data Analytics Software and Applications (Hadoop or Spark)	3	1	3/ Spring	CS102, CS203	Statistics
	STA326	Data Science Practice	3	1	3/ Spring	STA303/ CS303/ CS303B	Statistics
	Total		18				
Practice-based Courses	STA490	Thesis (Design)	12	12	4/ Spring		Statistics
	STA480	Research and innovation projects**	2	2	Any semester		Statistics

	STA470	Professional Internship**	2	2	Summer Vacation		Statistics
	Total		14	14			
Total			52	18			

Notes:
1. Students must choose one of the scientific research and innovation projects (including various scientific research activities, innovative projects in science and technology, winning competitions at the provincial level or above, publishing papers, further studies at home and abroad and attending a certain amount of seminars, etc., which will be recognized by the department for credit) and professional internship to carry out their practice. Students can choose to carry out their research and innovation projects and professional internship in any semester after the first academic year, with a minimum requirement of 4 weeks for professional internship.
2.Some courses are subject to change, so please take the corresponding courses in the actual semester of the course offering unit.

Table 2 Major Elective Courses

Course Code	Course Name	Credits	Practice-based Learning Credits	Terms	Prerequisite	Dept.
MA109/MA111/MA121	Linear algebra intensive/ Higher Algebra II/ Higher Algebra II (H)	4		1/Spring	MA113	Mathematics
CS307	Database Principles	3	1	2/ Autumn	CS102B	Computer Science
STA217	Introduction to Data Science	3		2/ Autumn	MA102a/MA127	Statistics
MA206	Mathematical modeling	3		2/ Spring	MA201a/MA230 /MA201b	Mathematics
MA201a/ MA230	The ordinary differential equation A/ Ordinary differential equation A(H)	3		2/ Spring	MA203a/MA213-16) and (MA109/ (MA111/MA121)	Mathematics
MA202/MA232	Functions of complex variables/ Function of complex variable (H)	3		2/ Spring	MA203a/MA213-16) and (MA109/MA111/ MA121)	Mathematics
MA208	Applying Stochastic Processes	3		2/ Spring	MA213-16 and (MA215/MA212) and (MA109/MA111/) (MA121)	Mathematics
MA309	Time Series Analysis	3		3/ Autumn	MA204/MA212	Statistics
STA320	Statistical Learning	3		3/ Spring	MA329	Statistics
MA305	Numerical analysis	3		3/ Autumn	MA203a/MA213-16	Mathematics
MA301	Real variable function*	3		3/ Autumn	MA203a/MA213-16	Mathematics
CS305	Computer Networks	3	1	3/ Autumn	CS109	Computer Science

MAT7035	Calculate statistics	3		3/ Autumn	MA204	Statistics
MA308	Statistical computing and software	3		3/ Autumn	MA329	Statistics
STA327	Generalized Linear Models	3		3/ Autumn	MA329	Statistics and Data Science
CS306	Data Mining	3	1	3/ Spring	CS203/CS203B	Computer Science
STA306	Bayesian statistics	3		3/ Spring	MA329	Statistics
MA417	Non-parametric statistics	3		3/ Spring	MA212/MA204	Statistics
CS324	Deep Learning	3		3/ Spring	CS303	Computer Science
MA409	Statistical Data Analysis (SAS)	3		3/ Spring	MA329	Statistics
STA404	Network Science and Computing	3		3/ Spring	MA212	Statistics
STA435	Statistical English Writing and Speaking	3		3-4/spring	None	Statistics
CS405	Machine Learning	3	1	4/ Autumn	MA113 and MA212	Computer Science
MA405	Survival Analysis	3		4/ Autumn	MA329	Statistics
Total		73	4			

Notes:

1、 Students enrolled in Mathematical Analysis I and II will receive 18 credits of elective courses, and students enrolled in Advanced Mathematics I and II will receive 20 credits of elective courses.

2 、 After this training program is developed, all new courses offered by the Department of Statistics and Data Science can be recognized as elective credits for the Data Science and Big Data Technology program.

3 、 Some courses may be offered in different semesters, so please take the courses in the actual semester of the course offering unit. The additional credits of mathematics and science courses in the science experimental classes can be used as credits of professional elective courses.

Table 3 Overview of Practice-based Courses

Course Code	Course Name	Credits	Practice-based Learning Credits	Terms	Prerequisite	Dept.
STA470	Professional Internship*	2	2	Summer Vacation		Statistics
STA480	Research and innovation projects*	2	2	Any semester		Statistics
STA490	Thesis (Design)	12	12	4/ Spring		Statistics
CS109	Fundamentals of Computer Programming	3	1	1-2/Spring	None	Compute r Science
MA110	MATLAB Programming	3	1	2/ Spring	None	Mathematics
MA206	Mathematical modeling	3		2/ Spring	MA201a/ MA230/MA201 b	Mathematics
CS205	C/C++ Programming	3	1	1/Spring	None	Computer Science
CS203 /CS203B	Data Structures and Algorithm Analysis/ Data Structures and Algorithm Analysis B	3	1	2/ Autumn	CS102A	Computer Science
CS332	Information Retrieval	3	1	3/ Spring	CS203	Computer Science
CS306	Data Mining	3	1	3/ Spring	CS203/CS203B	Computer Science
CS405	Machine Learning	3	1	4/ Autumn	MA113, MA212	Computer Science
MA308	Statistical computing and software	3	1	3/ Autumn	MA204/MA212	Statistics
Total		48	27			

Data Science and Big Data Technology Program Curriculum Structure

Time	First Year	Second Year	3rd grade	3rd/4th grade
Autumn	Mathematical Analysis I / Calculus I	Mathematical Analysis III / Mathematical Analysis	Artificial Intelligence B	Machine Learning
	Advanced Linear Algebra I / Linear Algebra	Discrete Mathematics	Distributed Storage and Parallel Computing	Survival Analysis
		Foundation of Probability Theory	Statistical linear model	Time Series Analysis
		Database Principles	Natural Language Processing	Numerical analysis
		Introduction to Data Science		Functions of real variables
				Calculate statistics
				Generalized Linear Models
				Statistical Computation and Software
				Computer Networks
Spring	Mathematical Analysis II / Calculus II	Mathematical Statistics	Big Data Analytics Software and its Applications	Graduation design
	Linear Algebra	Operations and optimization	Multivariate Statistical Analysis	Professional internships or research and innovation projects (Conducted in any semester in grades 2, 3 and 4)
		Data Structures and Algorithm Analysis B	Data Science Practice	Data Mining
		Applying Stochastic Processes	Statistical Data Analysis (SAS)	Nonparametric Statistics
		Ordinary differential equation A	Bayesian statistics	Deep Learning
		Functions of complex variables	Network Science and Computing	
		Mathematical Modeling	Statistical Writing and Communication in English	
			Statistical Learning	