

# BIostatISTICS CONCENTRATION (BIS)

The M.S. with a concentration in Biostatistics is a two-year program that provides training in clinical trials, epidemiologic methodology, implementation science, data science, statistical genetics, and mathematical models for infectious diseases. Students have a choice of three pathways: the *Biostatistics Standard Pathway*, the *Biostatistics Implementation and Prevention Science Methods Pathway*, and the *Biostatistics Data Science Pathway*. In contrast to the more general M.P.H. degree, the M.S. degree emphasizes the mastery of biostatistical skills from the beginning of the plan of study. While graduates of this program may apply to the Ph.D. degree program, the M.S. degree is itself quite marketable as a terminal degree. Part-time enrollment is permitted.

## DEGREE REQUIREMENTS

The Biostatistics concentration requires the completion of fifteen required and elective courses for the Standard Pathway and the Implementation and Prevention Sciences Pathway. Sixteen required and elective courses must be completed for the Data Science Pathway. These requirements exclude the Seminar, BIS 525/BIS 526; the Summer Internship, BIS 695; EPH 100; and EPH 101. *Note:* Half-term courses cannot count as an elective unless an additional half-term course is taken, and the Biostatistics faculty have approved both courses as an elective.

The Graduate School requires an overall grade average of High Pass, including grades of Honors in at least two full-term graduate courses for students enrolled in a two-year program. In order to maintain the minimum average of High Pass, each grade of Pass on the student's transcript must be balanced by one grade of Honors. Each grade of Fail must be balanced by two grades of Honors. If a student retakes a course in which the student has received a failing grade, only the newer grade will be considered in calculating this average. The initial grade of Fail, however, will remain on the student's transcript. A grade awarded at the conclusion of a full-year course in which no grade is awarded at the end of the first term would be counted twice in calculating this average.

## CURRICULUM

### Required Courses for All Pathways

(or substitutions approved by the student's adviser and the DGS)

BIS 525	Seminar in Biostatistics and Journal Club <sup>1</sup>	0
BIS 526	Seminar in Biostatistics and Journal Club <sup>1</sup>	0
BIS 623	Advanced Regression Models	1
or S&DS 612	Linear Models	
BIS 628	Longitudinal and Multilevel Data Analysis	1
BIS 630	Applied Survival Analysis	1
or BIS 643	Theory of Survival Analysis	
BIS 678	Statistical Practice I	1
BIS 695	Summer Internship in Biostatistics <sup>1</sup>	0

EPH 100	Professional Skills Series <sup>1</sup>	0
EPH 101	Professional Skills Series <sup>1</sup>	0
EPH 509	Fundamentals of Epidemiology	1
EPH 608	Frontiers of Public Health <sup>2</sup>	1
S&DS 541	Probability Theory <sup>3</sup>	1
or S&DS 551	Stochastic Processes	
or S&DS 600	Advanced Probability	
S&DS 542	Theory of Statistics <sup>3</sup>	1
or S&DS 610	Statistical Inference	

<sup>1</sup> These courses do not count toward the fifteen required courses.

<sup>2</sup> Students entering the program with an M.P.H. or relevant graduate degree may be exempt.

<sup>3</sup> These courses are offered in the Graduate School of Arts and Sciences.

### Additional Required Courses: Standard Pathway

BIS 679	Advanced Statistical Programming in SAS and R	1
BIS 681	Statistical Practice II	1

Two of the following suggested electives:

BIS 534	Stochastic Models and Inference for the Biomedical and Social Sciences	1
BIS 536	Measurement Error and Missing Data	1
BIS 537	Statistical Methods for Causal Inference	1
BIS 540	Fundamentals of Clinical Trials	1
BIS 550	Topics in Biomedical Informatics and Data Science	1
BIS 555	Machine Learning with Biomedical Data	1
BIS 560	Introduction to Health Informatics	1
BIS 567	Bayesian Statistics	1
BIS 568	Applied Machine Learning in Healthcare	1
BIS 620	Data Science Software Systems	1
BIS 629	Advanced Methods for Implementation and Prevention Science	1
BIS 631	Advanced Topics in Causal Inference Methods	1
BIS 633	Population and Public Health Informatics	1
BIS 634	Computational Methods for Informatics	1
BIS 638	Clinical Database Management Systems and Ontologies	1
BIS 640	User-Centered Design of Digital Health Tools	1
BIS 643	Theory of Survival Analysis <sup>1</sup>	1
BIS 645	Statistical Methods in Human Genetics	1
BIS 646	Nonparametric Statistical Methods and Their Applications	1
BIS 691	Theory of Generalized Linear Models	1

BIS 692	Statistical Methods in Computational Biology	1
CDE 566	Causal Inference Methods in Public Health Research	1
CDE 634	Advanced Applied Analytic Methods in Epidemiology and Public Health	1
EMD 553	Transmission Dynamic Models for Understanding Infectious Diseases	1
ENAS 912	Biomedical Image Processing and Analysis	1
HPM 573	Advanced Topics in Modeling Health Care Decisions	1
HPM 583	Methods in Health Services Research	1

Additional electives must be approved by the Standard Pathway director.

<sup>1</sup> Cannot fulfill elective credit if substituted for BIS 630.

Three electives in Statistics and Data Science. Suggested electives are:

CPSC 540	Database Design and Implementation <sup>1</sup>	1
CPSC 546	Data and Information Visualization <sup>1</sup>	1
CPSC 552	Deep Learning Theory and Applications <sup>1</sup>	1
CPSC 570	Artificial Intelligence <sup>1</sup>	1
CPSC 577	Natural Language Processing <sup>1</sup>	1
CPSC 582	Current Topics in Applied Machine Learning <sup>1</sup>	1
CPSC 583	Deep Learning on Graph-Structured Data <sup>1</sup>	1
CPSC 640	Topics in Numerical Computation <sup>1</sup>	1
CPSC 670	Topics in Natural Language Processing <sup>1</sup>	1
CPSC 677	Advanced Natural Language Processing <sup>1</sup>	1
CPSC 680	Trustworthy Deep Learning <sup>1</sup>	1
CPSC 752	Biomedical Data Science: Mining and Modeling <sup>1</sup>	1
INP 558	Computational Methods in Human Neuroscience <sup>1</sup>	1
INP 599	Statistics and Data Analysis in Neuroscience <sup>1</sup>	1
MGT 510	Data Analysis and Causal Inference <sup>3</sup>	2
MGT 556	Big Data & Customer Analytics <sup>3</sup>	2
MGT 803	Decision Making with Data <sup>3</sup>	2
S&DS 517	Applied Machine Learning and Causal Inference <sup>1</sup>	1
S&DS 530	Data Exploration and Analysis <sup>1</sup>	1
S&DS 551	Stochastic Processes <sup>1</sup>	1
S&DS 562	Computational Tools for Data Science <sup>1</sup>	1
S&DS 563	Multivariate Statistical Methods for the Social Sciences <sup>1</sup>	1
S&DS 565	Introductory Machine Learning <sup>1</sup>	1
S&DS 569	Numerical Linear Algebra: Deterministic and Randomized Algorithms <sup>1</sup>	1
S&DS 600	Advanced Probability <sup>1</sup>	1
S&DS 610	Statistical Inference <sup>1</sup>	1
S&DS 611	Selected Topics in Statistical Decision Theory <sup>1</sup>	1
S&DS 612	Linear Models <sup>1,2</sup>	1

S&DS 625	Statistical Case Studies <sup>1</sup>	1
S&DS 631	Optimization and Computation <sup>1</sup>	1
S&DS 632	Advanced Optimization Techniques <sup>1</sup>	1
S&DS 661	Data Analysis <sup>1</sup>	1
S&DS 662	Statistical Computing <sup>2</sup>	1
S&DS 663	Computational Mathematics Situational Awareness and Survival Skills <sup>1</sup>	1
S&DS 664	Information Theory <sup>1</sup>	1
S&DS 665	Intermediate Machine Learning <sup>1</sup>	1
S&DS 674	Applied Spatial Statistics <sup>1</sup>	1
S&DS 685	Theory of Reinforcement Learning <sup>1</sup>	1
Additional electives must be approved by the Standard Pathway director.		

<sup>1</sup> These courses are offered in the Graduate School of Arts and Sciences.

<sup>2</sup> Cannot fulfill elective credit if substituted for BIS 623.

<sup>3</sup> These courses are offered in the School of Management

Students wishing to complete a thesis may enroll in BIS 649 and BIS 650, Master's Thesis Research. This would be an additional requirement and cannot replace any of the required courses noted above. All students who complete a thesis will be required to present their research during a public seminar to the Biostatistics faculty and students in order to graduate.

*Additional Required Courses: Implementation and Prevention Science Methods Pathway*

BIS 629	Advanced Methods for Implementation and Prevention Science	1
BIS 679	Advanced Statistical Programming in SAS and R	1
BIS 681	Statistical Practice II <sup>1</sup>	1
EMD 533	Implementation Science	1

<sup>1</sup> A master's thesis is strongly recommended in place of BIS 681 and one elective. All students who complete a thesis will be required to present their research during a public seminar to the Biostatistics faculty and students in order to graduate.

At least one of the following:

BIS 536	Measurement Error and Missing Data	1
BIS 537	Statistical Methods for Causal Inference	1
BIS 631	Advanced Topics in Causal Inference Methods	1

At least two of the following:

CDE 516	Principles of Epidemiology II	1
CDE 534	Applied Analytic Methods in Epidemiology	1
EMD 538	Quantitative Methods for Infectious Disease Epidemiology	1
HPM 570	Cost-Effectiveness Analysis and Decision-Making <sup>1</sup>	1
HPM 575	Evaluation of Global Health Policies and Programs	1

HPM 586	Microeconomics for Health Policy and Health Management	1
HPM 587	Advanced Health Economics	1
SBS 541	Community Health Program Evaluation <sup>1</sup>	1
SBS 574	Developing a Health Promotion and Disease Prevention Intervention	1
SBS 580	Qualitative Research Methods in Public Health <sup>1</sup>	1
SBS 676	Questionnaire Development	1
S&DS 565	Introductory Machine Learning <sup>2</sup>	1
Alternative electives must be approved by the Implementation Science Pathway director.		

<sup>1</sup> These courses are highly recommended.

<sup>2</sup> This course is offered in the Graduate School of Arts and Sciences.

## Additional Required Courses: Data Science Pathway

BIS 620	Data Science Software Systems	1
BIS 687	Data Science Capstone	1

Two of the following Biostatistics, Computer Science or Statistical Methods courses

BIS 536	Measurement Error and Missing Data	1
BIS 537	Statistical Methods for Causal Inference	1
BIS 540	Fundamentals of Clinical Trials	1
BIS 550	Topics in Biomedical Informatics and Data Science	1
BIS 555	Machine Learning with Biomedical Data	1
BIS 567	Bayesian Statistics	1
BIS 629	Advanced Methods for Implementation and Prevention Science	1
BIS 634	Computational Methods for Informatics	1
BIS 645	Statistical Methods in Human Genetics	1
BIS 646	Nonparametric Statistical Methods and Their Applications	1
CB&B 752	Biomedical Data Science: Mining and Modeling <sup>1</sup>	1
CPSC 519	Full Stack Web Programming <sup>1</sup>	1
CPSC 526	Building Distributed Systems <sup>1</sup>	1
CPSC 539	Software Engineering <sup>1</sup>	1
CPSC 565	Theory of Distributed Systems <sup>1</sup>	1
CPSC 577	Natural Language Processing <sup>1</sup>	1
CPSC 640	Topics in Numerical Computation <sup>1</sup>	1
EMD 553	Transmission Dynamic Models for Understanding Infectious Diseases	1
S&DS 541	Probability Theory <sup>4</sup>	1
S&DS 551	Stochastic Processes <sup>5</sup>	1
S&DS 611	Selected Topics in Statistical Decision Theory <sup>1</sup>	1

S&DS 625	Statistical Case Studies	1
S&DS 661	Data Analysis <sup>1</sup>	1
S&DS 664	Information Theory	1
Additional electives must be approved by the Data Science Pathway director.		

One of the following Machine Learning courses:

BIS 555	Machine Learning with Biomedical Data <sup>2</sup>	1
BIS 568	Applied Machine Learning in Healthcare	1
BIS 634	Computational Methods for Informatics <sup>2</sup>	1
BIS 691	Theory of Generalized Linear Models	1
CB&B 555	Unsupervised Learning for Big Data <sup>1</sup>	1
CB&B 567	Topics in Deep Learning: Methods and Biomedical Applications <sup>1</sup>	1
CB&B 663	Deep Learning Theory and Applications <sup>1</sup>	1
CB&B 745	Advanced Topics in Machine Learning and Data Mining <sup>1</sup>	1
CPSC 569	Randomized Algorithms <sup>1</sup>	1
CPSC 583	Deep Learning on Graph-Structured Data	1
CPSC 644	Geometric and Topological Methods in Machine Learning	1
CPSC 670	Topics in Natural Language Processing <sup>1</sup>	1
S&DS 517	Applied Machine Learning and Causal Inference <sup>1</sup>	1
S&DS 538	Probability and Statistics	1
S&DS 562	Computational Tools for Data Science <sup>1</sup>	1
S&DS 565	Introductory Machine Learning <sup>1</sup>	1
S&DS 569	Numerical Linear Algebra: Deterministic and Randomized Algorithms	1
S&DS 631	Optimization and Computation <sup>1</sup>	1
S&DS 632	Advanced Optimization Techniques <sup>1</sup>	1
S&DS 665	Intermediate Machine Learning <sup>1</sup>	1
S&DS 674	Applied Spatial Statistics	1
S&DS 684	Statistical Inference on Graphs	1
S&DS 685	Theory of Reinforcement Learning	1
S&DS 686	High-Dimensional Phenomena in Statistics and Learning	1
Additional electives must be approved by the Data Science Pathway director.		

One of the following Database courses:

BIS 550	Topics in Biomedical Informatics and Data Science <sup>2</sup>	1
BIS 638	Clinical Database Management Systems and Ontologies	1
BIS 679	Advanced Statistical Programming in SAS and R	1
CPSC 537	Introduction to Database Systems <sup>1</sup>	1
MGT 660	Advanced Management of Software Development <sup>3</sup>	4
Additional electives must be approved by the Data Science Pathway director.		

<sup>1</sup> These courses are offered in the Graduate School of Arts and Sciences

- <sup>2</sup> These courses can only be counted to fulfill the requirement of one category; they cannot be counted twice
- <sup>3</sup> These courses are offered at the School of Management
- <sup>4</sup> Cannot fulfill elective if taken as a requirement
- <sup>5</sup> Cannot fulfill elective if taken as a substitute for S&DS 541

Two additional electives are required from the Biostatistics, Machine Learning, or Database list. Other courses from Public Health or other departments must be approved by the Data Science Pathway director.

Students wishing to complete a thesis may enroll in BIS 649 and BIS 650, Master's Thesis Research. This would be an additional requirement and cannot replace any of the required courses noted above. All students who complete a thesis will be required to present their research during a public seminar organized by the Biostatistics department.

## COMPETENCIES

Upon receiving an M.S. in the Biostatistics concentration of Public Health, the student will be able to:

- Select from a variety of analytical tools to test statistical hypotheses, interpret results of statistical analyses, and use these results to make relevant inferences from data.
- Design efficient computer programs for study management, statistical analysis, as well as presentation using R, SAS, and other programming languages.
- Demonstrate oral and written communication and presentation skills to effectively communicate and disseminate results to professional audiences.