Statistics

Degrees Offered

- Master of Science
- Master of Data Science
- Doctor of Philosophy

Nature of the Programs

The Department of Statistics offers a Master of Science (M.S.) in Statistics, a Master of Data Science (M.D.S.), and a Doctor of Philosophy (Ph.D.) in Computational Statistics. The department also offers a Certificate in Applied Statistics (C.A.S.). The M.S. degree is intended to qualify the student to assume a professional role in educational, industrial, or governmental research projects; to teach in a college; or to undertake advanced training toward a doctorate in statistics or one of the quantitative fields of science. The M.D.S degree is intended for professionals or students who want applied statistics and data science courses to enhance their job opportunities and quantitative knowledge. The Ph.D. degree is designed for students who are interested in the statistical and data sciences and who want to work in the following: 1) a research university, 2) collaboratively with researchers in the computational sciences at research institutes, 3) the government, or 4) research-driven corporations.

Because many students receive baccalaureate degrees from colleges that do not offer undergraduate programs in statistics, and because historically statistics has been primarily a field of graduate education, a student does not need a degree in statistics to enter the degree programs in statistics. A good background in mathematics, science, or engineering is reasonable preparation for graduate work in statistics.

The Department of Statistics also participates in the Combinatorial Computing and Discrete Mathematics (CCDM) Area of Emphasis within the Computer and Information Science Ph.D. Program or the Mathematics Ph.D. Program.

FACULTY

PROFESSORS

- Erdogan Gulel - Ph.D. (State University of New York, Buffalo)
  Bayesian Inference, Biostatistics, Categorical Data Analysis
- E. James Harner - Ph.D. (Cornell U.)
  Dynamic graphics, Statistical computing and modeling, Statistical education.
- Robert Mnatsakanov - Ph.D. (Moscow State Institute of Electronics and Mathematics)
  Nonparametric statistics, Statistical Inverse Problems, Mixture Models, Change-set Problems

ASSOCIATE PROFESSORS

- Mark V. Culp - Ph.D. (University of Michigan)
  Statistical Machine Learning, Computational Statistics, Semi-supervised and Multi-view Learning, Biometrics
- Gerald R. Hobbs Jr. - Ph.D. (Kansas State University)
  Biostatistics, Nonparametric Statistics, Regression Analysis
- Kenneth J. Ryan - Ph.D. (Iowa State University)
  Experimental Design, Statistical Machine Learning, Biometrics

ASSISTANT PROFESSOR

- Erin R. Leatherman - Ph.D. (Ohio State)
  Prediction and Design for Computer and Physical Experiments.
- Philip Turk - Ph.D. (Montana State University)
  Statistical Computing, Adaptive Cluster Sampling, Response Surface Methodology, Statistical Applications to Biological and Environmental Problems

CLINICAL ASSISTANT PROFESSOR

- Huey Miin Lee - Ph.D. (Johns Hopkins University)
  Bioinformatics, Statistical Education

RESEARCH ASSISTANT PROFESSOR

- Yanqing Hu - Ph.D. (University of Virginia)
  Design of Clinical Trials, Adaptive Designs, Nonparametric Statistics, Bioinformatics
INSTRUCTOR
• Sarah Quesen - M.P.H. (West Virginia University)
  Biostatistics, Health Policy, Statistical Education

RESEARCH ASSOCIATE
• Anthony A. Billings - M.S. (West Virginia University), A.B.D. (Carnegie Mellon University)
  Statistical Computing, Statistical Modeling, Robust Estimation, Nonlinear Dynamic Systems, Statistical Education

ADJUNCT PROFESSOR
• Michael Andrew - Ph.D. (University of Wyoming)
  Epidemiological Analysis, Statistical Modeling, Time Series Analysis

ADJUNCT ASSISTANT PROFESSOR
• Stacey Culp - Ph.D. (University of Michigan)
  Functional Data Analysis, Statistical Consulting
• Fekedulegn B. Desta - Ph.D.
  Categorical Data Analysis, Multivariate Statistical Methods, Nonlinear Regression, Forest Growth Modeling, SAS Programming

PROFESSOR EMERITUS
• William V. Thayne - Ph.D. (University of Illinois)
  Experimental Design, Statistical Genetics, Regression Analysis
• Edwin C. Townsend - Ph.D. (Cornell University)
  Experimental Design, Regression Analysis

ASSOCIATE PROFESSOR EMERITUS
• Daniel M. Chilko - M.S. (Rutgers University)
  Statistical Computing, Computer Graphics

ADMISSIONS AND PREREQUISITES FOR MASTER OF SCIENCE IN STATISTICS
Students are expected to know the material contained in the following courses or areas upon admission to the program. Otherwise, these deficiencies must be removed as early as possible in the student’s degree program under the terms specified by the Admissions and Standards Committee.

• Single and multivariable calculus (MATH 155, MATH 156, MATH 251, or equivalent)
• Linear or matrix algebra (MATH 441 or equivalent)
• Probability and statistics (STAT 215 or equivalent)
• Knowledge of a high-level programming language

ADMISSIONS AND PREREQUISITES FOR MASTER OF DATA SCIENCE
The Master of Data Science is designed for a Fall admission for full-time students. This will allow students to complete the MDS by the end of the summer in the following year through an intensive program of study, i.e., in 15 months. Students can lessen the course load during the academic year by entering the MDS in the Summer term. Students entering the Spring term may require an extra semester to complete the program due to scheduling constraints. Part-time students, including part-time online students, can enter the program at any time.

A West Virginia University graduate student can petition the Director to enter the program as a transfer student or as part of a dual major. However, transfer students and degree students within West Virginia University are required to take at least 12 of their 30 credits after admission into the Master of Data Science program. Financial aid is not given for this program by the Department of Statistics.

The prerequisite for admission is college algebra and single-variable calculus. Matrix algebra and knowledge of a high-level programming language are recommended, but are not required for admission. Students without single-variable calculus can be admitted provisionally, but this deficiency must be removed during their first or second term of enrollment, e.g., by taking MATH 150 or equivalent.

Beyond the above mathematical entrance requirements, the Director will base admission on the applicant’s:

• résumé;
• transcript;
• graduate (and/or undergraduate) GPA and major;
• statement of interest.
The GRE General Test is not required for admission.

To maintain major status and to graduate, students must maintain at least a 3.0 GPA in all courses counting towards the Master of Data Science. Students falling below a 3.0 GPA for a semester on courses counting towards this degree will be placed on academic probation and will have one year to raise their GPA to a 3.0.

ADMISSIONS AND PREREQUISITES FOR THE CERTIFICATE IN APPLIED STATISTICS

Students can apply for the Certificate in Applied Statistics during the time, or prior to the time, they are taking STAT 511 or STAT 512. Students who have taken courses beyond these two basic courses can petition the Academic Standards Committee for acceptance into the program.

The prerequisite for admission is a college algebra course. Single and multi-variable calculus are recommended but not required for admission. However, certain elective courses have calculus prerequisites. Applicants must have a baccalaureate degree.

Beyond the above mathematical entrance requirements, the Admissions and Standards Committee will base admission on the following material submitted by the applicant:

- Resume or curriculum vitae
- Transcript

The GRE General Test is not required for admission.

Master of Science in Statistics
Master of Data Science

MASTER OF SCIENCE IN STATISTICS

DEGREE REQUIREMENTS

To obtain a Master of Science in Statistics, the student must complete the course and comprehensive examination requirements. The student must maintain a minimum GPA of 3.0 and earn a grade of C or better in all courses counting towards the degree.

COURSE REQUIREMENTS

CORE COURSES

<table>
<thead>
<tr>
<th>Course</th>
<th>Title</th>
<th>Credit</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 512</td>
<td>Statistical Methods 2</td>
<td>3</td>
</tr>
<tr>
<td>STAT 513</td>
<td>Design of Experiments</td>
<td>3</td>
</tr>
<tr>
<td>STAT 545</td>
<td>Applied Regression Analysis</td>
<td>3</td>
</tr>
<tr>
<td>STAT 555</td>
<td>Categorical Data Analysis</td>
<td>3</td>
</tr>
<tr>
<td>STAT 561</td>
<td>Theory of Statistics 1</td>
<td>3</td>
</tr>
<tr>
<td>STAT 562</td>
<td>Theory of Statistics 2</td>
<td>3</td>
</tr>
</tbody>
</table>

TOTAL CREDIT HOUR REQUIREMENT

<table>
<thead>
<tr>
<th>Course</th>
<th>Title</th>
<th>Credit</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 582</td>
<td>Statistical Consulting</td>
<td>1</td>
</tr>
<tr>
<td>STAT 590</td>
<td>Teaching Practicum</td>
<td>1</td>
</tr>
<tr>
<td>STAT 696</td>
<td>Graduate Seminar</td>
<td>1</td>
</tr>
<tr>
<td>Electives</td>
<td></td>
<td>6 - 12</td>
</tr>
<tr>
<td>STAT 697</td>
<td>Research</td>
<td>1 - 6</td>
</tr>
</tbody>
</table>

**Non-thesis Option:** Thirty-three credit hours of Statistics courses at the 500**, 600, or 700 level, including the core courses, are required to complete an M.S. in Statistics without a thesis.

**Thesis Option:** Twenty-seven credit hours of Statistics courses at the 500**, 600, or 700 level, including the core courses, are required to complete an M.S. in Statistics with a thesis. In addition, the student must complete six credits of STAT 697 under the direction of a faculty advisor over two consecutive semesters, produce a written thesis, and orally defend the thesis before a faculty committee.

* Non-STAT electives require departmental consent.

** Credit towards the degree requirement is not given for STAT 511 Statistical Methods 1.
EXAMINATIONS
Students must pass a written comprehensive examination on foundational material. The examination covers the theory taught in STAT 561 and STAT 562 and the applications taught in STAT 512, STAT 513, and STAT 545. The exam is given twice a year on the Thursday during the second full week following spring semester final exams and on the third Saturday in October.

MASTER OF DATA SCIENCE
The Master of Data Sciences is designed for students and professionals who want data science courses to enhance their job opportunities and quantitative knowledge. This course-based master's degree includes an experiential component/capstone that integrates the practical application of statistical and data science skills with professional knowledge. This program is designed for off-campus delivery. However, students can take the program on-campus or by combining online and distance-based courses. Students can be full-time or part-time and schedule courses according to job-related or other constraints. Due to the nature of this program and the required skills for being a data scientist, the curriculum is largely prescribed.

This master's program will provide students a solid foundation in data analysis, data modeling, big data processing, predictive analytics, statistical computing, massive and streaming data mining, high-dimensional data analysis, high-performance analytics, and data/web technologies. Together, these skills will enhance the analytical abilities of graduate students and the quantitative skills of professionals.

DEGREE REQUIREMENTS
Students in the Master of Data Science program must complete 30 credit hours.

REQUIRED COURSES

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
<th>Credit Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSCI 501</td>
<td>Data Analysis</td>
<td>3</td>
</tr>
<tr>
<td>DSCI 502</td>
<td>Data Modeling</td>
<td>3</td>
</tr>
<tr>
<td>DSCI 503</td>
<td>Data Science</td>
<td>3</td>
</tr>
<tr>
<td>DSCI 504</td>
<td>Data Visualization</td>
<td>3</td>
</tr>
<tr>
<td>DSCI 601 or DSCI 602</td>
<td>High-Dimensional Data Analysis or Massive Data Mining</td>
<td>3</td>
</tr>
</tbody>
</table>

Total Hours 15

TOTAL CREDIT HOUR REQUIREMENT
Required Courses 15
Electives* 12
DSCI 682 or DSCI 689: 3
Total Hours 30

RECOMMENDED ELECTIVE COURSES

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
<th>Credit Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 521</td>
<td>SAS Programming</td>
<td>3</td>
</tr>
<tr>
<td>STAT 522</td>
<td>Advanced SAS Programming</td>
<td>3</td>
</tr>
<tr>
<td>STAT 523</td>
<td>Statistical Computing</td>
<td>3</td>
</tr>
<tr>
<td>DSCI 601 or DSCI 602</td>
<td>High-Dimensional Data Analysis or Massive Data Mining</td>
<td>3</td>
</tr>
<tr>
<td>STAT 623</td>
<td>Data Technologies</td>
<td>3</td>
</tr>
<tr>
<td>STAT 624</td>
<td>High Performance Analytics</td>
<td>3</td>
</tr>
</tbody>
</table>

EXPERIENTIAL COMPONENT/CAPSTONE
DSCI 682 (Data Science Practicum) or DSCI 689 (Professional Field Experience): The experiential component includes a capstone project integrating data science skills with professional knowledge. The capstone project is developed by the student in conjunction with faculty and, whenever possible, with employers; supervised by faculty and, whenever possible, with employers; and evaluated by faculty. Students should apply the methodologies learned in this program to a real-world problem in data science. As such, this course should be taken at or near the end of a student's program of study. A reproducible report on the methodologies used and the results is required for this experiential component/capstone.

* Electives without a DSCI or STAT subject code require departmental approval.

More information concerning masters-level graduate studies may be found at the Department of Statistics website (http://www.stat.wvu.edu).

Ph.D. in Computational Statistics
The Computational Statistics doctoral program is designed for students who are interested in the statistical and data sciences and who also want to work collaboratively with researchers in the computational sciences, e.g., computational biology, computational finance, and computational chemistry.
Students graduating in this major will be in strong demand by universities as teaching and research faculty; by high technology, biotechnology, financial, and pharmaceutical companies among others; and by various governmental agencies and research institutes.

Students enrolled in this major will become research statisticians, and they will develop deep theoretical and conceptual understanding of and strong skills in data technologies, statistical computation, high-performance analytics, data mining, statistical machine learning, statistical modeling, and bioinformatics. They will also have the option of pursuing a theoretical track. Students in this program will develop strong statistical consulting and grant-writing skills and be able to apply these skills to collaborative research projects, including the ability to write reproducible research-based documents.

ALL ADMITTED STUDENTS MUST COMPLETE THE FOLLOWING:

• Pass the qualifying exam by the end of their first year.
• Complete the core M.S. courses (listed below) within the first two years of study.

To maintain major status and to graduate, students must maintain at least a 3.0 GPA in all courses counting towards the Ph.D. Students falling below a 3.0 GPA for a semester on courses counting towards the Ph.D. will be placed on academic probation and will have one year to raise their GPA to a 3.0.

RESIDENCY REQUIREMENTS

Residency requires at least two consecutive semesters of full-time work (at least nine hours per term) after being admitted to the doctoral program. Courses taken for non-degree credit will not count for graduate residency.

PROGRAM REQUIREMENTS

Required M.S. Courses

<table>
<thead>
<tr>
<th>Course</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 512</td>
<td>Statistical Methods 2</td>
<td>3</td>
</tr>
<tr>
<td>STAT 513</td>
<td>Design of Experiments</td>
<td>3</td>
</tr>
<tr>
<td>STAT 523</td>
<td>Statistical Computing</td>
<td>3</td>
</tr>
<tr>
<td>STAT 541</td>
<td>Applied Multivariate Analysis</td>
<td>3</td>
</tr>
<tr>
<td>STAT 545</td>
<td>Applied Regression Analysis</td>
<td>3</td>
</tr>
<tr>
<td>STAT 555</td>
<td>Categorical Data Analysis</td>
<td>3</td>
</tr>
<tr>
<td>STAT 561</td>
<td>Theory of Statistics 1</td>
<td>3</td>
</tr>
<tr>
<td>STAT 562</td>
<td>Theory of Statistics 2</td>
<td>3</td>
</tr>
<tr>
<td>STAT 582</td>
<td>Statistical Consulting</td>
<td>1</td>
</tr>
<tr>
<td>STAT 682</td>
<td>Statistics Practicum</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>26</td>
</tr>
</tbody>
</table>

Students entering the Ph.D. program must have at least the equivalent of these courses. If not, courses representing deficient areas must be taken within the first two years of study.

QUALIFYING EXAMINATION

The Ph.D. qualifying exam consists of the theory and applied parts of the M.S. Comprehensive Exam. Students who have not passed this exam with a 70 on each part prior to admittance into the doctoral program must take the Qualifying Exam prior to the end of their first year, including the summer term. The theory and applied parts are given separately and students must obtain a 70 on each. Students falling short of passing are allowed to retake one or both parts as necessary the next time the exam is given.

REQUIRED PHD COURSES

Ph.D. candidates must complete at least forty-six semester credit hours, including at least twelve hours of research, beyond the required M.S.-level statistics core courses.

A minimum GPA of 3.0 is required in all courses and a minimum grade of C must be earned in all required STAT courses

<table>
<thead>
<tr>
<th>Course</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 623</td>
<td>Data Technologies</td>
<td>3</td>
</tr>
<tr>
<td>STAT 624</td>
<td>High Performance Analytics</td>
<td>3</td>
</tr>
<tr>
<td>STAT 723</td>
<td>Data Mining</td>
<td>3</td>
</tr>
<tr>
<td>STAT 745</td>
<td>Data Mining</td>
<td>3</td>
</tr>
<tr>
<td>STAT 746</td>
<td>Stochastic Processes</td>
<td>3</td>
</tr>
<tr>
<td>STAT 763</td>
<td>Stat Methods-Bioinformatics</td>
<td>3</td>
</tr>
<tr>
<td>or STAT 765</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STAT 782</td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>
At least twelve credit hours from STAT 500-, 600-, and 700-level normal graded courses not previously taken as part of the courses counted towards the Statistics M.S.

<table>
<thead>
<tr>
<th>Course</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 796</td>
<td>1</td>
</tr>
<tr>
<td>STAT 797</td>
<td>Research 12</td>
</tr>
</tbody>
</table>

Total Hours: 46

Students who have taken more than the required hours for the M.S. degree core courses can apply some or all of these extra courses to the Ph.D. as long as all of the other Ph.D. requirements are met. The student’s dissertation committee can also approve elective courses from other programs if these courses contribute to the student’s program of study.

**RECOMMENDED ELECTIVES**

Students contemplating an academic or a research-oriented career should take these advanced theoretical statistics courses: STAT 761 Theoretical Statistics 1 and STAT 762 Theoretical Statistics 2.

**RESEARCH**

Ph.D. candidates must take at least twelve hours of research and at least one hour of STAT 796, where the student’s research is presented.

**EXAMINATIONS**

Ph.D. students must pass two written comprehensive area exams with a score of 70 on each. These student can choose two from the following list:

- A computation exam based on STAT 623 Data Technologies, STAT 624 High Performance Analytics, and STAT 723 Computational Statistics
- A statistical machine learning exam based on STAT 745 Data Mining and STAT 746 Statistical Machine Learning
- A statistical theory exam based on STAT 761 Theoretical Statistics 1 and STAT 762 Theoretical Statistics 2

**DISSERTATION**

The student should form a dissertation committee after passing the Comprehensive Exam. The dissertation committee must consist of a chair, who must be a member of the graduate faculty, and at least three other committee members, a majority of whom must be on the graduate faculty. The dissertation committee approves the student’s final plan of study. The student must also develop a dissertation topic in conjunction with his or her major professor and present this topic to the committee. Once the dissertation committee approves the area of research, the student is directed to develop a prospectus. Once the student’s major professor approves the prospectus, the student is directed to schedule the Oral Candidacy Exam. The prospectus must be complete in the sense that the student presents a solid case that his or her area of research can be completed.

**ORAL CANDIDACY EXAM**

The student must present and defend the prospectus. The student becomes a Ph.D. candidate once this exam is passed. If the student is not passed, he or she may be given another opportunity to revise the prospectus or begin a new research area at the discretion of the dissertation committee.

**ORAL DEFENSE**

The student must present and defend the results of the research as proposed in the Oral Candidacy Examination. If the student is not passed, he or she may be given another opportunity to retake the Oral Defense after the dissertation is revised at the discretion of the dissertation committee.

Students must follow all ECAS and university requirements relating to graduate studies and research throughout their entire graduate program. This includes the Electronic Dissertation Program.

**Certificate in Applied Statistics**

The Certificate in Applied Statistics (C.A.S.) can be earned through traditional classroom delivery as well as distance-based delivery. Many of the courses are offered both on-campus and online, including the required courses and the SAS and Data Science tracks.

The C.A.S. is designed for professionals or students who want applied statistics and data science courses to enhance their job opportunities and quantitative skills. If prerequisite requirements are met, theoretical courses in statistics are available as an option. This certificate program is designed for both on and off-campus delivery.

This certificate program will provide students with a solid foundation in statistical methods, predictive analytics, statistical computing, and data technologies. This program will enhance the quantitative skills of professionals or the research productivity of graduate students.

The C.A.S. is based on a coherent group of courses representing a specialized area of knowledge. To meet this objective each student must choose a track reflecting his or her interests. Certain tracks, e.g., the SAS, Data Science, and Modeling tracks, are offered as on and off-campus options. Other tracks are residential, although certain courses within these tracks may be offered in a distance-based format.
To maintain their status and to receive a Certificate in Applied Statistics, students must maintain at least a 3.0 GPA in courses counted towards the certificate. Students falling below a 3.0 GPA for a semester on courses counting towards this degree will be notified, and they will have one year to raise their GPA to a 3.0.

**REQUIRED COURSES**

Students in the certificate program must complete at least fifteen credit hours. The Certificate in Applied Statistics should represent a distinct area of study, which builds on the foundational required statistics courses. The distinct areas of study are: SAS, data science, modeling, applied statistics, and mathematical statistics. The courses required for the completion of each of these areas are defined below.

<table>
<thead>
<tr>
<th>Course</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 511</td>
<td>Statistical Methods 1</td>
<td>3</td>
</tr>
<tr>
<td>STAT 512</td>
<td>Statistical Methods 2</td>
<td>3</td>
</tr>
<tr>
<td>Three courses from one of the tracks listed below</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Total Hours</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

**TRACK OPTIONS AND COURSES**

**SAS TRACK:**

<table>
<thead>
<tr>
<th>Course</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 521</td>
<td>SAS Programming</td>
<td>3</td>
</tr>
<tr>
<td>STAT 522</td>
<td>Advanced SAS Programming</td>
<td>3</td>
</tr>
<tr>
<td>STAT 540</td>
<td>Intro-Exploratory Data Analysis</td>
<td>3</td>
</tr>
</tbody>
</table>

**DATA SCIENCE TRACK:**

<table>
<thead>
<tr>
<th>Course</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 523</td>
<td>Statistical Computing</td>
<td>3</td>
</tr>
<tr>
<td>STAT 623</td>
<td>Data Technologies</td>
<td>3</td>
</tr>
<tr>
<td>STAT 624</td>
<td>High Performance Analytics</td>
<td>3</td>
</tr>
</tbody>
</table>

**MODELING TRACK:**

<table>
<thead>
<tr>
<th>Course</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 541</td>
<td>Applied Multivariate Analysis</td>
<td>3</td>
</tr>
<tr>
<td>STAT 545</td>
<td>Applied Regression Analysis</td>
<td>3</td>
</tr>
<tr>
<td>STAT 547</td>
<td>Survival Analysis</td>
<td>3</td>
</tr>
</tbody>
</table>

**APPLIED STATISTICS TRACK:**

<table>
<thead>
<tr>
<th>Course</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 513</td>
<td>Design of Experiments</td>
<td>3</td>
</tr>
<tr>
<td>STAT 531</td>
<td>Sampling Theory and Methods</td>
<td>3</td>
</tr>
<tr>
<td>STAT 551</td>
<td>Nonparametric Statistics</td>
<td>3</td>
</tr>
</tbody>
</table>

**MATHEMATICAL STATISTICS TRACK:**

<table>
<thead>
<tr>
<th>Course</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 516</td>
<td>Forensic Statistics</td>
<td>3</td>
</tr>
<tr>
<td>STAT 561</td>
<td>Theory of Statistics 1</td>
<td>3</td>
</tr>
<tr>
<td>STAT 562</td>
<td>Theory of Statistics 2</td>
<td>3</td>
</tr>
<tr>
<td>Total Hours</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

**CUSTOMIZED TRACK:**

A student can propose a customized track to the Admissions and Standards Committee for approval based on three cohesive courses selected from 500 or 600-level courses taught by the Department of Statistics or two courses taught by the Department of Statistics and one statistics course taught by another department.

**CAPSTONE EXPERIENCE**

The certificate capstone will test the student’s ability to apply statistical methods to a real dataset using statistical computing tools. The student is required to develop a reproducible report including a brief description of the methodologies used, the results, and the conclusions. The report is limited to ten pages. The dataset type and methods used will depend on the track the student has chosen.

The capstone experience will be administered by the student’s advisor. The student will be given five days to complete the report at a date mutually agreed upon by the student and his or her advisor at or near the end of the student’s coursework. The advisor will determine whether the student passes or fails based on an agreed upon rubric. A student who fails will receive written information on the reasons for failure and be given three additional days to improve his or her report to the standards required for passing.
COURSES

STAT 505. Foundations-Probability/Stat. 3 Hours.
PR: MATH 156 or consent. Probability, random variables, discrete and continuous probability distributions, point and interval estimation, chi-square tests, linear regression, and correlation.

STAT 511. Statistical Methods 1. 3 Hours.
PR: MATH 126. Statistical models, distributions, probability, random variables, tests of hypotheses, confidence intervals, regression, correlation, transformations, F and Chi-square distributions, analysis of variance and multiple comparisons. (Equivalent to EDP 613 and PSYC 511.).

STAT 512. Statistical Methods 2. 3 Hours.
PR: STAT 511 or equivalent. Completely random, randomized complete block, Latin square, and split-plot experimental designs. Unplanned and planned multiple and orthogonal comparisons for qualitative and quantitative treatments and factorial arrangements. Multiple linear regression and covariance analysis. (Equivalent to EDP 614 and PSYC 512.).

STAT 513. Design of Experiments. 3 Hours.
PR: STAT 512 or equivalent. Expected mean squares, power of tests and relative efficiency for various experimental designs. Fixed, random, and mixed models. Use of sub-sampling, covariance, and confounding to increase power and efficiency.

STAT 516. Forensic Statistics. 3 Hours.
PR: STAT 215 or equivalent. Probabilistic and statistical evaluation of evidence in forensic science: concepts of uncertainty variation, discriminating power, coincidence/significance probabilities, historical overview, transfer evidence, DNA profiling, fingerprint identification, biometric identification, and case studies.

STAT 521. SAS Programming. 3 Hours.

STAT 522. Advanced SAS Programming. 3 Hours.
PR: STAT 521 or consent. Advanced topics in Statistical Analysis System (SAS); SAS SQL to generate reports, join tables, construct queries; SAS Macrolanguage basics; write/implement SAS macro programs. Prepares students for SAS Advanced Programmer Certification Exam.

STAT 523. Statistical Computing. 3 Hours.
PR: STAT 512. Monte Carlo methods; randomization, partitioning, and the bootstrap; identifying data structures, estimating functions, including density functions; statistical models of dependencies. R programming.

STAT 525. Statistical Graphics. 3 Hours.
PR: STAT 512. Introduction to R graphics; traditional graphs; the grid graphics model; lattice graphics; developing new graphics functions and objects in R. Visualizing large datasets.

STAT 531. Sampling Theory and Methods. 3 Hours.
PR: STAT 511 or consent. Survey components, methods of sampling for finite and infinite populations, single and multi-stage procedures, confidence limits for estimating population parameters, sample size determination, area sampling sources of survey error, and basic inference derived from survey design.

STAT 540. Intro-Exploratry Data Analysis. 3 Hours.
PR: An introductory statistics course. Basic ways in which observations given in counted and measured form are approached. Pictorial and arithmetic techniques of display and discovery. Methods employed are robust, graphical, and informal. Applications to social and natural sciences. (Alternate years.).

STAT 541. Applied Multivariate Analysis. 3 Hours.
PR: STAT 511 or equivalent. Introduction to Euclidean geometry and matrix algebra; multiple and multivariate regression including multiple and canonical correlation; the k-sample problem including discriminant and canonical analysis; and structuring data by factor analysis, cluster analysis, and multi-dimensional scaling.

STAT 543. Bioinformatics Data Analysis. 3 Hours.
PR: STAT 512 or equivalent. Statistical analyses of high-throughput experiments using data visualization, clustering, multiple testing, classification and other unsupervised and supervised learning methods. Data processing, including background adjustment and normalization. Case studies.

STAT 545. Applied Regression Analysis. 3 Hours.
PR: STAT 512 or equivalent. Matrix approach to linear and multiple regression, selecting the “best” regression equation, model building, and the linear models approach to analysis of variance and analysis of covariance.

STAT 547. Survival Analysis. 3 Hours.
PR: STAT 512. Survival model methodology, including model selection for incomplete data with censored, truncated, and interval censored observations. Applications to many real life problems using R.

STAT 551. Nonparametric Statistics. 3 Hours.
PR: STAT 511 or equivalent. Distribution-free procedures of statistical inference. Location and scale tests for homogeneity with two or more samples (related or independent); tests against general alternatives. (Alternate years.).
STAT 555. Categorical Data Analysis. 3 Hours.
PR: STAT 512 or equivalent. Bivariate association for ordinal and nominal variables, models for categorical or continuous responses as a special case of generalized linear models, methods for repeated measurement data, exact small-sample procedures.

STAT 561. Theory of Statistics 1. 3 Hours.
PR: MATH 251. Probability and random variables, univariate and multivariate distributions, expectations, generating functions, marginal and conditional distributions, independence, correlation, functions of random variables, including order statistics, limiting distributions, and stochastic convergence.

STAT 562. Theory of Statistics 2. 3 Hours.
PR: STAT 561. Techniques of point and interval estimation; properties of estimates including bias, consistency, efficiency, and sufficiency; hypothesis testing including likelihood ratio tests and Neyman-Pearson Lemma; Bayesian procedures; analysis of variance and nonparametrics.

STAT 582. Statistical Consulting. 1 Hour.
PR: STAT 513 or Consent. Statistical consulting principles and procedures. The entire consulting experience, including design, models, communication skills, ethics, tracking, and documentation, is presented in a series of case studies, including student presentations and reports on assigned cases.

STAT 590. Teaching Practicum. 1-3 Hours.
PR: Consent. Supervised practice in college teaching of statistics. Note: This course is intended to ensure that graduate assistants are adequately prepared and supervised when they are given college teaching responsibility. It will also present a mechanism for students not on assistantships to gain teaching experience. (Grading may be S/U.).

STAT 591A-Z. Advanced Topics. 1-6 Hours.
PR: Consent. Investigation in advanced topics not covered in regularly scheduled courses.

STAT 593A-Z. Special Topics. 1-6 Hours.
A study of contemporary topics selected from recent developments in the field.

STAT 595. Independent Study. 1-6 Hours.
STAT 595. Independent Study. 1-6 HR. Faculty supervised study of topics not available through regular course offerings.

STAT 623. Data Technologies. 3 Hours.
PR: STAT 512 or consent. R data manipulation and processing. Topics include: R operators, functions, data structures, and objects; R data input and output, package development, and text processing; R interfaces to XML and SQL databases.

STAT 624. High Performance Analytics. 3 Hours.
PR: STAT 623. High performance and data-stream computing using R. Topics include: parallel R packages; Hadoop clusters; MapReduce R scripting; shared R network spaces; beyond-memory data analysis; data-stream modeling and visualization.

STAT 641. Multivariate Statistical Thry. 3 Hours.
PR: STAT 541, and STAT 561 or consent. Euclidean vector space theory and matrix algebra, multivariate normal sampling theory, the theory of the multivariate general linear hypothesis including multivariate regression, MANOVA, and MANCOVA, and the theory of factor analysis.

STAT 645. Linear Models. 3 Hours.
PR: STAT 545 and STAT 362 or consent. Multivariate normal distribution, distribution of quadratic forms, linear models, general linear hypotheses, experimental design models, components of variance for random effects models.

STAT 682. Statistics Practicum. 1 Hour.
PR: STAT 582. Statistical consulting on university-related research projects under the direction of a statistics faculty member.

STAT 689. Professional Field Experience. 1-6 Hours.
PR: Consent. (May be repeated up to a maximum of 18 hours). Prearranged experiential learning program, to be planned, supervised, and evaluated for credit by faculty and field supervisors. Involves temporary placement with public or private enterprise for professional competence development.

STAT 690. Teaching Practicum. 1-3 Hours.
PR: Consent. Supervised practice in college teaching of statistics. Note: This course is intended to ensure that graduate assistants are adequately prepared and supervised when they are given college teaching responsibility. It also provides a mechanism for students not on assistantships to gain teaching experience. (Grading may be S/U.).

STAT 691A-Z. Advanced Topics. 0-6 Hours.
PR: Consent. Investigation of advanced topics not covered in regularly scheduled courses.

STAT 692A-Z. Directed Study. 1-6 Hours.
Directed study, reading, and/or research.

STAT 693A-Z. Special Topics. 1-6 Hours.
A study of contemporary topics selected from recent developments in the field.

STAT 694A-Z. Seminar. 1-6 Hours.
Special seminars arranged for advanced graduate students.

STAT 695. Independent Study. 1-6 Hours.
Faculty supervised study of topics not available through regular course offerings.
STAT 696. Graduate Seminar. 1 Hour.
PR: Consent. Each graduate student will present at least one seminar to the assembled faculty and graduate student body of his or her program.

STAT 697. Research. 1-15 Hours.
PR: Consent. Research activities leading to thesis, problem report, research paper or equivalent scholarly project, or a dissertation. (Grading may be S/U).

STAT 698. Thesis. 1-6 Hours.
PR: Consent. This is an optional course for programs that wish to provide formal supervision during the writing of student reports (698), or dissertations (798). Grading is normal.

STAT 699. Graduate Colloquium. 1-6 Hours.
PR: Consent. For graduate students not seeking coursework credit but who wish to meet residency requirements, use of the University’s facilities, and participate in its academic and cultural programs. Note: Graduate students who are not actively involved in coursework or research are entitled, through enrollment in their department’s 699/799 Graduate Colloquium to consult with graduate faculty, participate in both formal and informal academic activities sponsored by their program, and retain all of the rights and privileges of duly enrolled students. Grading is P/F; colloquium credit may not be counted against credit requirements for masters programs. Registration for one credit of 699/799 graduate colloquium satisfies the University requirement in the semester in which graduation occurs.

STAT 745. Data Mining. 3 Hours.
PR: STAT 545 or equivalent. Development of predictive models for large datasets, including logistic and linear models, regression and classification trees, and neural networks. Data preparation, including imputation and filtering.

STAT 761. Theoretical Statistics 1. 3 Hours.
PR: STAT 562 or consent. Advanced statistical theory including: consistent estimators; limiting distributions; asymptotic properties; goodness-of-fit tests; maximum likelihood estimation, moment generating functions; properties of statistical tests and procedures for finite-dimensional and infinite-dimensional spaces.

STAT 762. Theoretical Statistics 2. 3 Hours.
PR: STAT 761. Continuation of STAT 761 including: asymptotic optimality, contiguity of probability measures, local asymptotic normality of likelihood ratio test, Bayesian estimation, general linear models estimation and testing, and kernel smoothing methods in density and regression estimation.

STAT 763. Stochastic Processes. 3 Hours.
PR: STAT 561. Modeling of random phenomenon occurring over time, space, or time and space simultaneously. Modern techniques, such as the martingale decomposition, are applied to different statistical models.

STAT 765. Stat Methods-Bioinformatics. 3 Hours.

STAT 791A-Z. Advanced Topics. 1-6 Hours.
PR: Consent. Investigation of advanced topics not covered in regularly scheduled courses.

STAT 797. Research. 1-15 Hours.
PR: Consent. Research activities leading to thesis, problem report, research paper or equivalent scholarly project, or a dissertation. (Grading will be S/U).