DATA SCIENCE, MASTER’S DEGREE

Program Requirements
All master’s degrees in the Department of Applied Mathematics and Statistics ordinarily require a minimum of two semesters of registration as a full-time resident graduate student.

- Online Data Ethics course: Students must take an approved online data ethics course.
- Complete training on the responsible and ethical conduct of research. Please see WSE Policy on the Responsible Conduct of Research (https://engineering.jhu.edu/wse-research/resources-policies-forms/responsible-conduct-of-research/)
- Complete training on academic ethics.

A course grade of B- or better is required to meet all course requirements.

<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Core Requirements</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EN.553.636</td>
<td>Introduction to Data Science</td>
<td>4.0</td>
</tr>
</tbody>
</table>

**Core Areas**
Select one course in each of the four Core Areas: 12-16

**Statistics**
- EN.553.630 Introduction to Statistics
- EN.553.632 Bayesian Statistics
- EN.553.730 Statistical Theory
- EN.553.731 Statistical Theory II

**Machine Learning**
- PH.140.644 Statistical Machine Learning: Methods, Theory, and Applications
- EN.553.740 Machine Learning I
- EN.553.743 (EN.553.743 Graphical Models)

**Optimization**
- EN.553.761 Nonlinear Optimization I
- EN.553.762 Nonlinear Optimization II
- EN.553.765 Convex Optimization

**Computing**
- EN.553.688 Computing for Applied Mathematics
- EN.601.620 Parallel Programming

**Focus Areas**
Select three courses in one of the following focus areas: 9-12

**Computational Medicine**
- AS.410.633 Introduction to Bioinformatics
- AS.410.635 Bioinformatics: Tools for Genome Analysis
- AS.410.671 Gene Expression Data Analysis and Visualization
- EN.553.650 Computational Molecular Medicine
- EN.605.620 Algorithms for Bioinformatics
  or EN.605.626 Foundations of Algorithms
- EN.605.653 Computational Genomics
- EN.605.754 Analysis of Gene Expression and High-Content Biological Data

**Mathematics of Data Science**
- EN.553.633 Monte Carlo Methods
- EN.553.665 Introduction to Convexity
- EN.553.738 High-Dimensional Approximation, Probability, and Statistical Learning
- EN.553.740 Machine Learning I
- EN.553.761 Nonlinear Optimization I
- EN.553.762 Nonlinear Optimization II
- EN.553.763 Stochastic Search & Optimization
- EN.553.765 Convex Optimization
- EN.553.766 Combinatorial Optimization
- EN.553.792 Matrix Analysis and Linear Algebra
- EN.601.634 Randomized and Big Data Algorithms
- EN.601.635 Approximation Algorithms

**Language and Speech**
- AS.050.617 Semantics I
- AS.050.622 Semantics II

**Computational Machine Learning**
- PH.140.644 Statistical Machine Learning: Methods, Theory, and Applications
- EN.520.612 Machine Learning for Signal Processing
- EN.520.647 Information Theory
- EN.520.648 Compressed Sensing and Sparse Recovery
- EN.520.651 Random Signal Analysis
- EN.553.740 Machine Learning I
- EN.553.743 (EN.553.743 Graphical Models)
- EN.601.675 Machine Learning
- EN.601.676 Machine Learning: Data to Models
- EN.601.681 Machine Learning: Optimization
- EN.601.677 Causal Inference
- EN.601.679 (EN.601.679 Machine Learning: Representation Learning)
- EN.601.775 Statistical Machine Learning

**Computer Vision**
- EN.520.648 Compressed Sensing and Sparse Recovery
- EN.601.661 Computer Vision
- EN.601.682 Machine Learning: Deep Learning
- EN.601.783 Vision as Bayesian Inference

**Computational Finance**
- EN.553.627 Stochastic Processes and Applications to Finance
- EN.553.628 Stochastic Processes and Applications to Finance II
- EN.553.641 Equity Markets and Quantitative Trading
- EN.553.642 Investment Science
- EN.553.644 Introduction to Financial Derivatives
- EN.553.645 Interest Rate and Credit Derivatives
- EN.553.646 Risk Measurement/Management in Financial Markets
- EN.553.647 Quantitative Portfolio Theory and Performance Analysis
- EN.553.648 Financial Engineering and Structured Products
- EN.553.649 Advanced Equity Derivatives
- EN.553.753 Commodity Markets and Trade Finance
<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS.050.620</td>
<td>Syntax I</td>
</tr>
<tr>
<td>AS.050.625</td>
<td>Phonology I</td>
</tr>
<tr>
<td>EN.520.666</td>
<td>Information Extraction</td>
</tr>
<tr>
<td>EN.520.680</td>
<td>Speech and Auditory Processing by Humans and Machines</td>
</tr>
<tr>
<td>EN.601.665</td>
<td>Natural Language Processing</td>
</tr>
<tr>
<td>EN.601.765</td>
<td>Machine Learning: Linguistic &amp; Sequence Modeling</td>
</tr>
</tbody>
</table>

**Statistical Theory**

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>PH.140.644</td>
<td>Statistical Machine Learning: Methods, Theory, and Applications</td>
</tr>
<tr>
<td>EN.553.632</td>
<td>Bayesian Statistics</td>
</tr>
<tr>
<td>EN.553.730</td>
<td>Statistical Theory</td>
</tr>
<tr>
<td>EN.553.731</td>
<td>Statistical Theory II</td>
</tr>
<tr>
<td>EN.553.735</td>
<td>Topics in Statistical Pattern Recognition</td>
</tr>
<tr>
<td>EN.553.737</td>
<td>Distribution-free statistics and Resampling Methods</td>
</tr>
<tr>
<td>EN.553.738</td>
<td>High-Dimensional Approximation, Probability, and Statistical Learning</td>
</tr>
<tr>
<td>EN.553.739</td>
<td>Statistical Pattern Recognition Theory &amp; Methods</td>
</tr>
<tr>
<td>EN.601.677</td>
<td>Causal Inference</td>
</tr>
<tr>
<td>EN.601.775</td>
<td>Statistical Machine Learning</td>
</tr>
</tbody>
</table>

**Elective**

The program requires the student to take one elective course. To maximize a student’s flexibility in choosing this course, the student may choose any course offered at JHU that is directly or indirectly related to data science. The elective course must be approved by the student’s advisor as well as the Internal Oversight Committee.

**Capstone Experience**

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>EN.553.806</td>
<td>Capstone Experience in Data Science (EN.553.806 Capstone Experience)</td>
</tr>
</tbody>
</table>

Total Credits: 34-48

1 Or another project-oriented course approved by the faculty advisor and the Internal Oversight Committee. In addition to taking the capstone course, the student must write a paper on a related topic, approved in advance by the course instructor. The content of the paper should include a deeper study of the pre-approved topic that allows the students to apply data analysis techniques learned in the program, and possibly to extend those ideas to more general settings or to new application areas. This paper will be summarized in a poster presented in a poster session organized at the end of each semester.