

Statistics for Data Science

Statistics for Data Science is about using statistical concepts and methods to extract and make sense of information contained in data. That information may be used for prediction, decision-making or understanding underlying phenomena, whether it is public health, finance, environmental studies, or technology.

The Statistics for Data Science Graduate Certificate from NJIT will provide an introduction to randomness, sampling, data generation, statistical modeling, data analysis and statistical computing, which participants can leverage to enter the fields of machine learning, data analytics and data science.

Who would be suited to take this program?

This program is suitable for people with a quantitative background and/or some experience with dealing with data, and who would like to obtain some grounding in statistical methods and concepts in the context of data science. Examples include actuaries who are seeking to add data science and statistical learning tools to the other quantitative methods they are already employing in their work; computer scientists who want to complement their computing skills with statistical training; people who are exploring transitioning into careers in data science.

The certificate program is designed to be flexible, requiring only one core course, allowing students to select three elective courses in order to tailor the certificate to their needs.

What are the prerequisites?

Some experience with programming, equivalent to CS100. A year of calculus, equivalent to Math 111 and 112. Although not required, Math 226 (Discrete Analysis) or equivalent is highly recommended.

What will I learn?

- **Math 661: Applied Statistics** – An introduction to basic foundational ideas in probability, statistical inference and data exploration and analysis methods.
- **Math 644: Regression Analysis Methods** – An in-depth study of linear regression, the most commonly used data analysis method. This course covers both the theoretical and applied aspects of regression
- **Math 660: Introduction to Statistical Computing with SAS and R** – An introduction to the use of the R and SAS statistical software for data analysis and simulation.
- **Math 678: Statistical Methods for Data Science** – An introduction to statistical learning in data science, covering topics like regularization, K-means, regression and classification trees, and cross-validation.
- **Math 664: Methods for Statistical Consulting** – This course covers applied statistical methods for data analysis, with case studies, examples in R, and group projects.
- **CS 675: Machine Learning** – This course covers implementation of machine learning methods in Python as well as some machine learning theory.
- **CS 644: Introduction to Big Data** – This course covers topics in big data from data generation, storage, management, to analytics, with a focus on computing solutions in high-performance networks.

Suggested program of study

- Math 661, Math 644, Math 660 and Math 678 – this is the recommended sequence appropriate for most students.
- Math 661, Math 644, Math 678 and CS 675 – for students who already have ample experience with the R programming language.
- Math 661, Math 678, Math 664 and CS 675 – for students who already have ample knowledge of regression analysis and the R programming language.

Why Study Statistical Learning at NJIT?

NJIT is situated in Newark, minutes from Newark Penn Station. Jersey City and New York City are also a short train ride away, providing easy access to these commercial areas with many companies that employ statisticians and data scientists. NJIT is a top research university, with faculty performing cutting-edge research and publishing in leading journals. NJIT also consistently ranks highly on added-value and diversity. Finally, the certificate provides balanced training on theoretical and applied aspects of statistical learning.

Into what industries might holders of this program find employment?

Pharmaceuticals, insurance, public health, technology, banking, in fact, any industry that can improve its business through better use of its data.

Prerequisites

Applicants should have a bachelor's degree with one year of calculus, and some experience with programming.

Related Degree Programs

All courses in this program are related to the NJIT MS in Data Science (especially the Statistics track).

Take Note

Check the course descriptions for more information. Some courses have prerequisites and must be taken in order.

Faculty Advisor: Ji Meng Loh/Andrew Pole

What are the Required Courses?

| Code | Title | Credits |
|--|--|----------------|
| Core Courses (3) | | |
| MATH 661 | Applied Statistics | 3 |
| Electives | | |
| Select three (3) courses of the following: | | 3 |
| MATH 644 | Regression Analysis Methods | 3 |
| MATH 660 | Introduction to statistical Computing with SAS and R * | 3 |
| or CS 636 | Data Analytics with R Program | |
| MATH 678 | Statistical Methods in Data Science | 3 |
| MATH 664 | Methods for Statistical Consulting | 3 |
| CS 675 | Machine Learning * | 3 |
| CS 644 | Introduction to Big Data * | 3 |

* Only a maximum of 2 CS electives are allowed.