

Data Science Major

www.Statistics.Pitt.edu

www.Mathematics.Pitt.edu

Revised: 08/2024

The rapidly expanding collection of massive amounts of data is leading to transformations across broad segments of industry, science, and society. These changes have sparked great demand for individuals with skills in managing and analyzing complex data sets. Such skills are interdisciplinary, involving ideas typically associated with computing, information processing, mathematics, and statistics as well as the development of new methodologies spanning these fields. The undergraduate Data Science major establishes a framework to train University of Pittsburgh students to participate fully in this data revolution.

This major will allow students to gain skill sets that span key areas of statistics, computer science and mathematics, with foundational training providing literacy in four areas (data, algorithmic, mathematical, and statistical) that every student needs to master data science; the development of expertise that connects theory to the solution of real-world problems; and specialization towards more specific career focuses. Completing this major will prepare students to work as a data science professional or to pursue graduate study in a direction involving data in a significant way.

Required courses for the Data Science major

The Data Science major requires the completion of 61 credits distributed as follows.

Foundational courses

All of the following courses are required except as noted.

CMPINF 0401 Intermediate Programming

CS 0445 Data Structures

MATH 0220 Analytic Geometry and Calculus 1

MATH 0230 Analytic Geometry and Calculus 2

MATH 0280 Introduction to Matrices **or**

MATH 1180 Linear Algebra or MATH 1185 Honors

MATH 0480 Applied Discrete Mathematics **or**

CS 0441 Applied Discrete Mathematics

STAT 1060 Data Science Foundations **or**

CMPINF 1061 Data Science Foundations

STAT 1151 Introduction to Probability **or**

STAT 1631 Intermediate Probability

STAT 1152 Introduction to Mathematical Statistics **or**

STAT 1632 Intermediate Mathematical Statistics

Expertise courses

All of the following courses are required except as noted.

CS 0590 Social Implications for Computing Technology

CS 1501 Algorithm Implementation

CS 1656 Introduction to Data Science

MATH 1101 Optimization

STAT 1261 Principles of Data Science

STAT 1361 Statistical Learning and Data Science **or**

CS 1675 Introduction to Machine Learning

Specialization

Students will have the opportunity to pursue an area of specialization through the selection of elective courses in one of the targeted directions – Data Analytics, Modeling, Computer Systems, or Data in Context. Students should select three courses from the same category for optimum focus. However, students may choose courses from across the categories to suit their interests.

Data Analytics

Students pursuing a data analytics specialization will enhance their ability to make sound inferences and decisions using the science and art of learning from data: specifically, the design, collection, analysis, and interpretation of data in an uncertain world, and the communication of findings.

CS 1671 Human Language Technologies

CS 1674 Introduction to Computer Vision

INFSCI 1068 GIS

INFSCI 1520 Information Visualization

INFSCI 1530 Data Mining

INFSCI 1560 Information Retrieval

MATH 1080 Numerical Linear Algebra

STAT 1201 Applied Nonparametric Statistics

STAT 1211 Applied Categorical Data Analysis

STAT 1221 Applied Regression

STAT 1231 Applied Experimental Design

STAT 1241 Applied Sampling

STAT 1311 Applied Multivariate Statistics

STAT 1321 Applied Time Series

STAT 1341 Sports Analytics

Modeling

Students pursuing a modeling specialization will enhance their ability to develop and harness theoretical tools to characterize structure within data and to represent and analyze processes that may underlie this structure.

CS 1538 Introduction to Simulation
MATH 1025 Introduction to Mathematical Cryptography
MATH 1070 Numerical Mathematical Analysis
MATH 1080 Numerical Linear Algebra
MATH 1310 Graph Theory
MATH 1360 Modeling in Applied Math
MATH 1510 Mathematical Theory of Probability
STAT 1211 Applied Categorical Data Analysis
STAT 1321 Applied Time Series
STAT 1731 Stochastic Processes
STAT 1741 Applied Probability

Computer Systems

Students pursuing this specialization will gain depth of knowledge in the development, deployment, and analysis of the complex computer and information systems necessary for tackling large-scale data science problems

CS 0449 Introduction to Systems Software
CS 1550 Introduction to Operating Systems
CS 1555 Databases **or**
INFSCI 1500 Databases
CS 1635 User Interfaces **or**
INFSCI 2415 User Interfaces
CS 1645 Introduction to High Performance Computing Systems
CS 2450 Parallel Computing
INFSCI 2125 Network Analysis

Capstone course

Select a capstone course, relevant to the chosen specialization, from the following list.

CMPINF 1981 Project Studio
MATH 1103 BIG Problems
STAT 1961 Statistical Data Science in Action
Any faculty-guided research project, subject to approval by the Data Science program director(s).

Grade requirements

A grade of C or better is required in each course that is to count toward the major. A minimum GPA of 2.0 in departmental courses is required for graduation.

Satisfactory/No Credit option

No course that counts toward the major can be taken on an S/NC basis.

Writing (W) requirement

CS 0590 fulfills the writing-intensive course for this major.

Advising

Lucas Mentch Department of Statistics LKM31@Pitt.edu	Jon Rubin Department of Mathematics JonRubin@Pitt.edu
Junshu Bao Department of Statistics JUB69@Pitt.edu	Armin Schikorra (MATH) Department of Mathematics Schikorra@Pitt.edu
Jason DeBlois Department of Mathematics Jdeblois@pitt.edu	David Swigon (MATH) Department of Mathematics Swigon@Pitt.edu

Checklist for the Data Science major

Foundational courses

_____ CS 0401
_____ CS 0445
_____ MATH 0220
_____ MATH 0230
_____ MATH 0280 or MATH 1180 or MATH 1185 Honors
_____ MATH 0480 or CS 0441
_____ STAT 1060 or CMPINF 1061
_____ STAT 1151 or STAT 1631
_____ STAT 1152 or STAT 1632

Expertise courses

_____ CS 0590
_____ CS 1501
_____ CS 1656
_____ MATH 1101
_____ STAT 1261
_____ STAT 1361 or CS 1675

Specialization courses

Capstone course
