

### **3.3.3.4 Data Science and Analytics**

Data science is an emerging field of study that involves statistical and computational principles, methods and systems for extracting and structuring knowledge from data. On a daily basis, large data sets are routinely generated by activities in the sciences, administration, leisure and commerce. Data scientists are constantly seeking patterns and predicting outcomes from these vast collections of data.

The four-year direct Honours programme in Data Science and Analytics (DSA) is designed to prepare graduates who are ready to acquire, manage and explore data that will inspire changes around the world. Singapore is a financial hub, with key industries focusing on biomedical sciences, health care, manufacturing, e-commerce and sustainable energy, among others. The DSA programme will equip its graduates with the skills to contribute to the activities of these industries. They will be able to handle problems like uncovering hidden stock market indicators, extracting information from medical images, predicting consumer behaviour and a host of similar interesting questions.

#### **Programme Structure and Curriculum Rationale**

The DSA programme is jointly offered by the Department of Mathematics and the Department of Statistics and Applied Probability in the Faculty of Science, with the collaboration of the School of Computing.

Students will read modules in Mathematics, Statistics and Computer Science, and be exposed to the interplay between these three key areas in the practice of data science. In their third and fourth years of study, students will also delve deep into subject matters such as computation and optimisation, computer algorithms, database and data processing, data mining and machine learning, and high-dimensional statistics. Students will also undertake an industry-driven capstone project module, where they will work with real-life data, providing them with an opportunity to tackle real-life issues and problems in a workplace environment.

#### **Co-Operative Education**

The NUS Co-Operative (Co-Op) Education Programme formally integrates academic studies with relevant work experience, where students complete multiple internship stints alternating with regular academic semesters over their candidature at NUS thus forming an integral part of the student's learning experience.

Students in the DSA programme have the option to participate in co-op education which comprises the following study/internship sequence:

	Semester 1	Semester 2	Special Term
Year 1	Study	Study	Study
Year 2	Study	Study	Internship (full time)
Year 3	Study & Internship (full time)	Internship (full time)	Internship (full time)
Year 4	Study & Internship (full time)	Study	

The first three internship segments ride on the Undergraduate Professional Internship Programme (UPIP) of the Faculty of Science. The last two internship segments take the form of an Honours-level project (DSA4299).

### **Career Prospects**

As the need for extensive data collection, processing and analyses increases across various sectors, DSA graduates can expect to build a career as data science professionals in both public and private firms, in industries ranging from technology to infocomm, transportation, telecommunications, e-commerce, etc.

### **Graduation Requirements**

To be awarded a B.Sc. or B.Sc. (Hons.) with a primary major in Data Science and Analytics, candidates must satisfy the following:

<b>Module Level</b>	<b>Major Requirements</b>	<b>Cumulative Major MCs</b>
<b>Level 1000 (16 MCs)</b>	Pass - CS1010/CS1010S/CS1010X Programming Methodology - DSA1101 Introduction to Data Science - MA1101R Linear Algebra I - MA1102R Calculus	16

<p><b>Level 2000 (24 MCs)</b></p>	<p>Pass</p> <ul style="list-style-type: none"> <li>- CS2040 Data Structures and Algorithms</li> <li>- DSA2101 Essential Data Analytics Tools: Data Visualisation</li> <li>- DSA2102 Essential Data Analytics Tools: Numerical Computation</li> <li>- MA2311 Techniques in Advanced Calculus or MA2104 Multivariable Calculus</li> <li>- ST2131/MA2216 Probability</li> <li>- ST2132 Mathematical Statistics</li> </ul>	<p>40</p>
<p><b>Levels 3000 and 4000 (56 MCs)</b></p>	<p>Pass</p> <ul style="list-style-type: none"> <li>- CS3244 Machine Learning</li> <li>- DSA3101 Data Science in Practice</li> <li>- DSA3102 Essential Data Analytics Tools: Convex Optimisation</li> <li>- ST3131 Regression Analysis</li> <li>- DSA4199 Honours Project in Data Science <i>or</i> DSA4299 Applied Project in Data Science</li> <li>- Six additional modules from List A and List B subject to the following restrictions: <ul style="list-style-type: none"> <li>+ There must be at least two modules each from List A and from List B1/ List B2</li> <li>+ A maximum of two DSA426x series modules can be used to fulfil this requirement</li> <li>+ There must be at least four modules at level 4000</li> </ul> </li> </ul>	<p>96</p>

**List A – DSA modules**

DSA4211 High-Dimensional Statistical Analysis

DSA4212 Optimisation for Large-Scale Data-Driven Inference

DSA426x Sense-making Case Analysis: YY and ZZ

**List B1 – DSA-recognised modules (no hidden pre-requisites)**

MA3236 Nonlinear Programming

MA3252 Linear and Network Optimisation

MA4270 Data Modelling and Computation

ST3232 Design and Analysis of Experiments

ST3233 Applied Time Series Analysis

ST3239 Survey Methodology

ST3240 Multivariate Statistical Analysis

ST3247 Simulation

ST3248 Statistical Learning I

ST4231 Computer Intensive Statistical Methods

ST4234 Bayesian Statistics

ST4248 Statistical Learning II

**List B2 – DSA-recognised modules (with hidden pre-requisites) \***

CS3210 Parallel Computing

CS3223 Database Systems Implementation

CS3230 Design and Analysis of Algorithms

CS3243 Introduction to Artificial Intelligence

CS4224 Distributed Databases

CS4225 Big Data Systems for Data Science or Massive Data Processing Techniques in Data Science

CS4231 Parallel and Distributed Algorithms

CS4234 Optimisation Algorithms

CS4243 Computer Vision and Pattern Recognition

CS4248 Natural Language Processing

CS5340 Uncertainty Modelling in AI

MA4230 Matrix Computation

\* Note: For List B2, i.e., the DSA-recognised modules with hidden pre-requisites, DSA students who wish to read these modules will be provided with academic advice by the Faculty/Department on their study plans where necessary, as such students would have to read 'additional' pre-requisite modules.

<b>Summary of Requirements</b>	<b>B.Sc. (Hons.)</b>
University Requirements	20 MCs
Faculty Requirements	8 MCs*
Major Requirements	96 MCs
Unrestricted Elective Modules	36 MCs
Total	160 MCs

\* 8 MCs of Faculty requirements are fulfilled through the reading of a CS-coded module and a ST/MA-coded module within the DSA curriculum.

Students are required to fulfill the remaining 8 MCs of Faculty requirements from any two of the following subject groups: Chemical Sciences, Life Sciences, Physical Sciences or Multidisciplinary & Interdisciplinary Sciences; but not from the following subject groups: Computing Sciences and Mathematical & Statistical Sciences.