FACULTY OF SCIENCE
THE UNIVERSITY OF HONG KONG
香港大學理學院

SHKU Science

Master of Data Science

Apply now for entry in **September 2023**

Machine Learning

Advanced Statistical Modelling

Data Mining

Cloud Computing

Spatial Data Analysis

In the era of big data, colossal amount of data is being generated every day from channels including social media, consumer profile, global economy, etc. These complications of data require us not only to describe and understand the information, but more importantly, to reveal crucial information to facilitate the best course of action. While traditional data processing application software is incapacitated, the launch of Master of Data Science (MDASC) provides students with the chance to acquire the quantitative and analytical skills much needed to stay ahead of the game.

Programme Information

Master of Data Science (MDASC) is a taught master programme jointly offered by Department of Statistics and Actuarial Science (host) and Department of Computer Science.

Its interdisciplinarity promotes the applications of computer technology, operational research, statistical modelling, and simulation to decision-making and problem-solving in all organizations and enterprises within the private and public sectors.

The curriculum of the MDASC programme adopts a well-balanced and comprehensive pedagogy of both statistical and computational concepts and methodologies, underpinning applications that are not limited to business or a single field alone.

Programme Highlights

- Interdisciplinary and comprehensive curriculum
- Solid foundation in statistical and computational analyses
- Electives cover a broad range of contemporary topics about Computer Science and Statistics
- Hands-on applications of methodologies with powerful software
- Capstone project with real-life scenario

Course Highlights

The core courses of the proposed MDASC programme mainly focus on both predictive and prescriptive concepts and methodologies with an effort to equip students with a solid foundation in statistical and computational analyses, e.g.

- Computational intelligence
- Time series forecasting
- Deep learning
- Statistical modelling

The electives cover a broad range of contemporary topics and provide students with solid training in diverse and applied techniques used in data science, including but not limited to:

- Financial data analysis
- Multimedia technologies
- Natural language processing
- Blockchain data analytics

Targeted Taught Postgraduate Programmes Fellowships Scheme

MDASC is selected as an eligible programme under the University Grants Committee for Targeted Taught Postgraduate Programmes Fellowships Scheme. Selected local students admitted to the MDASC (full-time or part-time) in the academic year 2023-24 are eligible to apply (with terms and conditions apply).

Local offer recipients who wish to apply for the Fellowship Scheme should prepare a proposal on how they can contribute to the priority areas (i.e. Research and STEM) of Hong Kong after completing the programme. Successful Fellowship Scheme applicants will each receive an award of HK$120,000.

Reimbursable Course(s) by Continuing Education Fund (CEF)*

持續進修基金
Continuing Education Fund

The following courses have been included in the list of reimbursable courses under the CEF:

- COMP7503 Multimedia technologies
- COMP7506 Smart phone apps development
- COMP7507 Visualization and visual analytics
- COMP7906 Introduction to cyber security
- STAT8017 Data mining techniques
- STAT8019 Marketing analytics

All CEF applicants are required to attend at least 70% of the courses before they are eligible for fee reimbursement.

*The mother programme (Master of Data Science) of these courses is recognised under the Qualifications Framework (QF Level 6).

"The COVID-19 pandemic served to accelerate the shift of consumers from physical stores to web-and-app based options. With millions of customers using this products, Statisticians, Mathematicians, and Data Scientists with the know-how to interpret the trends of shoppers are in ever-growing demand."

The 2021 Jobs Rated Report by www.careercast.com

Programme Curriculum

Commencing in September, the curriculum is composed of 72 credits of courses in either 1.5 academic years for full-time study, or 2.5 academic years for part-time study. Courses with 6 credits are offered in the first and second semesters while courses with 3 credits are normally offered in the summer semester. If a student selects a course whose contents are similar to a course (or courses) which he/she has taken in his/her previous study, the Department may not approve the selection in question. The curriculum is the same for both full-time and part-time study modes.

Compulsory Courses (36 credits)

COMP7404	Computational intelligence and machine learning (6 credits)
DASC7011	Statistical inference for data science (6 credits)
DASC7104	Advanced database systems (6 credits)
DASC7606	Deep learning (6 credits)
STAT7102	Advanced statistical modelling (6 credits)
STAT8003	Time series forecasting (6 credits)

Disciplinary Electives (24 credits)*

with at least 12 credits from List A and at least 12 credits from List B

List A

COMP7105	Advanced topics in data science (6 credits)
COMP7305	Cluster and cloud computing (6 credits)
COMP7409	Machine learning in trading and finance (6 credits)
COMP7503	Multimedia technologies (6 credits)
COMP7506	Smart phone apps development (6 credits)
COMP7507	Visualization and visual analytics (6 credits)
COMP7906	Introduction to cyber security (6 credits)
FITE7410	Financial fraud analytics (6 credits)
ICOM6044	Data science for business (6 credits)

List B

STAT6008	Advanced statistical inference (6 credits)
STAT6013	Financial data analysis (6 credits)
STAT6015	Advanced quantitative risk management (6 credits)
STAT6016	Spatial data analysis (6 credits)
STAT6019	Current topics in statistics (6 credits)
STAT7008	Programming for data science (6 credits)
STAT8017	Data mining techniques (6 credits)
STAT8019	Marketing analytics (6 credits)
STAT8306	Statistical methods for network data (3 credits)
STAT8307	Natural language processing and text analytics (3 credits)
STAT8308	Blockchain data analytics (3 credits) NEW

Capstone Requirement (12 credits)

DASC7600	Data science project (12 credits)

Remarks:

1. ***Students who have completed the same courses in their previous studies in HKU, e.g. Master of Statistics or Master of Science in Computer Science may, on production of relevant transcripts, be permitted to select up to 24 credits of disciplinary electives from either List A or List B above if they are not able to find any untaken options from either of the lists of disciplinary electives.**
2. **The programme structure will be reviewed from time to time and is subject to change.**

Course Description

Compulsory Courses

COMP7404 Computational intelligence and machine learning (6 credits)

This course will teach a broad set of principles and tools that will provide the mathematical, algorithmic and philosophical framework for tackling problems using Artificial Intelligence (AI) and Machine Learning (ML). AI and ML are highly interdisciplinary fields with impact in different applications, such as, biology, robotics, language, economics, and computer science. AI is the science and engineering of making intelligent machines, especially intelligent computer programs, while ML refers to the changes in systems that perform tasks associated with AI. Ethical issues in advanced AI and how to prevent learning algorithms from acquiring morally undesirable biases will be covered.

Topics may include a subset of the following: problem solving by search, heuristic (informed) search, constraint satisfaction, games, knowledge-based agents, supervised learning, unsupervised learning; learning theory, reinforcement learning and adaptive control and ethical challenges of AI and ML.

Pre-requisites: Nil, but knowledge of data structures and algorithms, probability, linear algebra, and programming would be an advantage.

Assessment: coursework (50%) and examination (50%)

DASC7011 Statistical inference for data science (6 credits)

Computing power has revolutionized the theory and practice of statistical inference. Reciprocally, novel statistical inference procedures are becoming an integral part of data science. By focusing on the interplay between statistical inference and methodologies for data science, this course reviews the main concepts underpinning classical statistical inference, studies computer-intensive methods for conducting statistical inference, and examines important issues concerning statistical inference drawn upon modern learning technologies. Contents include classical frequentist and Bayesian inferences, computer-intensive methods such as the EM algorithm, the bootstrap and the Markov chain Monte Carlo, large-scale hypothesis testing, high-dimensional modeling, and post-model-selection inference.

Assessment: coursework (40%) and examination (60%)

DASC7104 Advanced database systems (6 credits)

The course will study some advanced topics and techniques in database systems, with a focus on the aspects of database systems design & algorithms and big data processing for structured data. Traditional topics include: query optimization, physical database design, transaction management, crash recovery, parallel databases. This course will also survey some recent developments in selected areas such as NoSQL databases and SQL-based big data management systems for relational (structured) data.

Assessment: coursework (50%) and examination (50%)

DASC7606 Deep learning (6 credits)

Machine learning is a fast growing field in computer science and deep learning is the cutting edge technology that enables machines to learn from large-scale and complex datasets. Ethical implications of deep learning and its applications will be covered and the course will focus on how deep neural networks are applied to solve a wide range of problems in areas such as natural language processing, and image processing. Other applications such as financial predictions, game playing and robotics may also be covered. Topics covered include linear and logistic regression, artificial neural networks and how to train them, recurrent neural networks, convolutional neural networks, generative models, deep reinforcement learning and unsupervised feature learning.

Prerequisites: Basic programming skills, e.g., Python is required.

Assessment: coursework (40%) and examination (60%)

STAT7102 Advanced statistical modelling (6 credits)

This course introduces modern methods for constructing and evaluating statistical models and their implementation using popular computing software, such as R or Python. It will cover both the underlying principles of each modelling approach and the model estimation procedures. Topics from: (i) Linear regression models; (ii) Generalized linear models; (iii) Model selection and regularization; (iv) Kernel and local polynomial regression; selection of smoothing parameters; (v) Generalized additive models; (vi) Hidden Markov models and Bayesian networks.

Assessment: coursework (50%) and examination (50%)

STAT8003 Time series forecasting (6 credits)

A time series consists of a set of observations on a random variable taken over time. Such series arise naturally in climatology, economics, finance, environmental research and many other disciplines. In additional to statistical modelling, the course deals with the prediction of future behaviour of these time series. This course distinguishes different types of time series, investigates various representations for them and studies the relative merits of different forecasting procedures.

Assessment: coursework (40%) and examination (60%)

Disciplinary Electives

COMP7105 Advanced topics in data science (6 credits)

This course will introduce selected advanced computational methods and apply them to problems in data analysis and relevant applications.

Assessment: coursework (50%) and examination (50%)

COMP7305 Cluster and cloud computing (6 credits)

This course offers an overview of current cloud technologies, and discusses various issues in the design and implementation of cloud systems. Topics include cloud delivery models (SaaS, PaaS, and IaaS) with motivating examples from Google, Amazon, and Microsoft; virtualization techniques implemented in Xen, KVM, VMWare, and Docker; distributed file systems, such as Hadoop file system; MapReduce and Spark programming models for large-scale data analysis, networking techniques in hyper-scale data centers. The students will learn the use of Amazon EC2 to deploy applications on cloud, and implement a SPARK application on a Xen-enabled PC cluster as part of their term project.

Prerequisites: Students are expected to install various open-source cloud software in their Linux cluster, and exercise the system configuration and administration. Basic understanding of Linux operating system and some programming experiences (C/C++, Java or Python) in a Linux environment are required.

Assessment: coursework (50%) and examination (50%)

COMP7409 Machine learning in trading and finance (6 credits)

The course introduces our students to the field of Machine Learning, and help them develop skills of applying Machine Learning, or more precisely, applying supervised learning, unsupervised learning and reinforcement learning to solve problems in Trading and Finance.

This course will cover the following topics. (1) Overview of Machine Learning and Artificial Intelligence, (2) Supervised Learning, Unsupervised Learning and Reinforcement Learning, (3) Major algorithms for Supervised Learning and Unsupervised Learning with applications to Trading and Finance, (4) Basic algorithms for Reinforcement Learning with applications to optimal trading, asset management, and portfolio optimization, (5) Advanced methods of Reinforcement Learning with applications to high-frequency trading, cryptocurrency trading and peer-to-peer lending.

Assessment: coursework (65%) and examination (35%)

COMP7503 Multimedia technologies (6 credits)

This course presents fundamental concepts and emerging technologies for multimedia computing. Students are expected to learn how to develop various kinds of media communication, presentation, and manipulation techniques. At the end of course, students should acquire proper skill set to utilize, integrate and synchronize different information and data from media sources for building specific multimedia applications. Topics include media data acquisition methods and techniques; nature of perceptually encoded information; processing and manipulation of media data; multimedia content organization and analysis; trending technologies for future multimedia computing.

Assessment: coursework (50%) and examination (50%)

COMP7506 Smart phone apps development (6 credits)

Smart phones have become an essential part of our everyday lives. The number of smart phone users worldwide today surpasses six billion and is forecast to further grow by more than one billion in the next few years. Smart phones play an important role in mobile communication and applications.

Smart phones are powerful as they support a wide range of applications (called apps). Most of the time, smart phone users just download their favorite apps remotely from the app stores. There is a great potential for software developer to reach worldwide users.

This course aims at introducing the design and technical issues of smart phone apps. For example, smart phone screens are usually smaller than computer monitors while smart phones usually possess more hardware sensors than conventional computers. We have to pay special attention to these aspects in order to develop attractive and successful apps. Various modern smart phone apps development environments and programming techniques (such as Java for Android phones and Swift for iPhones) will also be introduced to facilitate students to develop their own apps.

Students should have basic programming knowledge.

Mutually exclusive with: COMP3330 Interactive mobile application design and programming

Assessment: coursework (60%) and examination (40%)

COMP7507 Visualization and visual analytics (6 credits)

This course introduces the basic principles and techniques in visualization and visual analytics, and their applications. Topics include human visual perception; color; visualization techniques for spatial, geospatial and multivariate data, graphs and networks; text and document visualization; scientific visualization; interaction and visual analysis.

Assessment: coursework (50%) and examination (50%)

COMP7906 Introduction to cyber security (6 credits)

The aim of the course is to introduce different methods of protecting information and data in the cyber world, including the privacy issue. Topics include introduction to security; cyber attacks and threats; cryptographic algorithms and applications; network security and infrastructure.

Mutually exclusive with: ICOM6045 Fundamentals of e-commerce security

Assessment: coursework (50%) and examination (50%)

FITE7410 Financial fraud analytics (6 credits)

This course aims at introducing various analytics techniques to fight against financial fraud. These analytics techniques include, descriptive analytics, predictive analytics, and social network learning. Various data set will also be introduced, including labeled or unlabeled data sets, and social network data set. Students learn the fraud patterns through applying the analytics techniques in financial frauds, such as, insurance fraud, credit card fraud, etc.

Key topics include: Handling of raw data sets for fraud detection; Applications of descriptive analytics, predictive analytics and social network analytics to construct fraud detection models; Financial Fraud Analytics challenges and issues when applied in business context.

Required to have basic knowledge about statistics concepts.

Assessment: coursework (60%) and examination (40%)

ICOM6044 Data science for business (6 credits)

The emerging discipline of data science combines statistical methods with computer science to solve problems in applied areas. In this case we focus on how data science can be used to solve business problems especially those in electronic commerce. By its very nature e-commerce is able to generate large amounts of data and data mining methods are quite helpful for managers in turning this data into knowledge which in turn can be used to make better decisions. These data sets and their accompanying quantitative methods have the potential to dramatically change decision making in many areas of business. For example, ideas like interactive marketing, customer relationship management, and database marketing are pushing companies to utilize the information they collect about their customers in order to make better marketing decisions.

This course focuses on how data science methods can be applied to solve managerial problems in marketing and electronic commerce. Our emphasis is developing a core set of principles that embody data science: empirical reasoning, exploratory and visual analysis, and predictive modeling. We use these core principles to understand many methods used in data mining and machine learning. Our strategy in this course is to survey several popular techniques and understand how they map into these core principles. These techniques are illustrated with case studies. However, the emphasis is not on the software for implementing these techniques but on understanding the inputs and outputs of these techniques and how they are used to solve business problems.

Assessment: coursework (65%) and examination (35%)

STAT6008 Advanced statistical inference (6 credits)

This course covers the advanced theory of point estimation, interval estimation and hypothesis testing. Using a mathematically-oriented approach, the course provides a formal treatment of inferential problems, statistical methodologies and their underlying theory. It is suitable in particular for students intending to further their studies or to develop a career in statistical research. Contents include: (1) Decision problem – frequentist approach: loss function; risk; decision rule; admissibility; minimaxity; unbiasedness; Bayes' rule; (2) Decision problem – Bayesian approach: prior and posterior distributions, Bayesian inference; (3) Estimation theory: exponential families; likelihood; sufficiency; minimal sufficiency; completeness; UMVU estimators; information inequality; large-sample theory of maximum likelihood estimation; (4) Hypothesis testing: uniformly most powerful (UMP) test; monotone likelihood ratio; UMP unbiased test; conditional test; large-sample theory of likelihood ratio; confidence set; (5) Nonparametric inference; bootstrap methods.

Assessment: coursework (40%) and examination (60%)

STAT6013 Financial data analysis (6 credits)

This course aims at introducing statistical methodologies in analyzing financial data. Financial applications and statistical methodologies are intertwined in all lectures. Contents include: recent advances in modern portfolio theory, copula, market microstructure, stochastic volatility models and high frequency data analysis.

Assessment: coursework (40%) and examination (60%)

STAT6015 Advanced quantitative risk management (6 credits)

This course covers statistical methods and models of risk management, especially of Value-at-Risk (VaR). Contents include: Value-at-risk (VaR) and Expected Shortfall (ES); univariate models (normal model, log-normal model and stochastic process model) for VaR and ES; models for portfolio VaR; time series models for VaR; extreme value approach to VaR; back-testing and stress testing.

Assessment: coursework (40%) and examination (60%)

STAT6016 Spatial data analysis (6 credits)

This course covers statistical concepts and tools involved in modelling data which are correlated in space. Applications can be found in many fields including epidemiology and public health, environmental sciences and ecology, economics and others. Covered topics include: (1) Outline of three types of spatial data: point-level (geostatistical), areal (lattice), and spatial point process. (2) Model-based geostatistics: covariance functions and the variogram; spatial trends and directional effects; intrinsic models; estimation by curve fitting or by maximum likelihood; spatial prediction by least squares, by simple and ordinary kriging, by trans-Gaussian kriging. (3) Areal data models: introduction to Markov random fields; conditional, intrinsic, and simultaneous autoregressive (CAR, IAR, and SAR) models. (4) Hierarchical modelling for univariate spatial response data, including Bayesian kriging and lattice modelling. (5) Introduction to simple spatial point processes and spatio-temporal models. Real data analysis examples will be provided with dedicated R packages such as geoR.

Assessment: coursework (50%) and examination (50%)

STAT6019 Current topics in statistics (6 credits)

This course includes two modules.

The first module, Causal Inference, is an introduction to key concepts and methods for causal inference. Contents include 1) the counterfactual outcome, randomized experiment, observational study; 2) Effect modification, mediation and interaction; 3) Causal graphs; 4) Confounding, selection bias, measurement error and random variability; 5) Inverse probability weighting and the marginal structural models; 6) Outcome regression and the propensity score; 7) The standardization and the parametric g-formula; 8) G-estimation and the structural nested model; 9) Instrumental variable method; 10) Machine learning methods for causal inference; 11) Other topics as determined by the instructor.

The second module, Functional data analysis, covers topics from: 1) Base functions; 2) Least squares estimation; 3) Constrained functions; 4) Functional PCA; 5) Regularized PCA; 6) Functional linear model; 7) Other topics as determined by the instructor.

Assessment: coursework (100%)

STAT7008 Programming for data science (6 credits)

In the big data era, it is very easy to collect huge amounts of data. Capturing and exploiting the important information contained within such datasets poses a number of statistical challenges. This course aims to provide students with a strong foundation in computing skills necessary to use Python to tackle some of these challenges. Possible topics to be covered may include exploratory data analysis and visualization, collecting data from a variety of sources (e.g. Excel, web-scraping, APIs and others), object-oriented programming concepts and scientific computation tools. Students will learn to create their own Python libraries.

Assessment: coursework (100%)

STAT8017 Data mining techniques (6 credits)

With the rapid developments in computer and data storage technologies, the fundamental paradigms of classical data analysis are mature for change. Data mining techniques aim at helping people to work smarter by revealing underlying structure and relationships in large amounts of data. This course takes a practical approach to introduce the new generation of data mining techniques and show how to use them to make better decisions. Topics include data preparation, feature selection, association rules, decision trees, bagging, random forests and gradient boosting, cluster analysis, neural networks, introduction to text mining.

Assessment: coursework (100%)

STAT8019 Marketing analytics (6 credits)

This course aims to introduce various statistical models and methodology used in marketing research. Special emphasis will be put on marketing analytics and statistical techniques for marketing decision making including market segmentation, market response models, consumer preference analysis and conjoint analysis. Contents include market response models, statistical methods for segmentation, targeting and positioning, statistical methods for new product design.

Assessment: coursework (40%) and examination (60%)

STAT8306 Statistical methods for network data (3 credits)

The six degree of separation theorizes that human interactions could be easily represented in the form of a network. Examples of networks include router networks, the World Wide Web, social networks (e.g. Facebook or Twitter), genetic interaction networks and various collaboration networks (e.g. movie actor coloration network and scientific paper collaboration network). Despite the diversity in the nature of sources, the networks exhibit some common properties. For example, both the spread of disease in a population and the spread of rumors in a social network are in sub-logarithmic time. This course aims at discussing the common properties of real networks and the recent development of statistical network models. Topics may include common network measures, community detection in graphs, preferential attachment random network models, exponential random graph models, models based on random point processes and the hidden network discovery on a set of dependent random variables.

Assessment: coursework (100%)

STAT8307 Natural language processing and text analytics (3 credits)

The textual data constitutes an enormous proportion of unstructured data which is characterized as one of 'V's in Big Data. The logical and computational reasonings are applied to transform large collection of written resources to structured data for use in further analysis, visualization, integration with structured data in database or warehouse, and further refinement using machine learning systems. This course introduces the methodology of text analytics. Topics include natural language processing, word representation, text categorization and clustering, topic modelling and sentiment analysis. Students are required to possess basic understanding of Python language.

Pre-requisites: Pass in STAT8017 Data mining techniques and DASC7606 Deep learning or equivalent

Assessment: coursework (100%)

STAT8308 Blockchain data analytics (3 credits) NEW

In this course, we start by studying the basic architecture of a blockchain. Then we move on to several major applications including (but not limited to) cryptocurrencies, fintech and smart contracts. We conclude by examining the cybersecurity issues facing the blockchain ecosystems.

Assessment: coursework (100%)

Capstone Requirement

DASC7600 Data science project (12 credits)

Candidate will be required to carry out independent work on a major project under the supervision of individual staff member. A written report is required.

Assessment: written report (75%) and oral presentation (25%)

Programme Duration

The programme normally extends over 1.5 academic years for full-time study, and 2.5 academic years for part-time study. Teaching will take place mostly on weekday evenings, and Saturday mornings and afternoons. All lectures are conducted in English at HKU.

Optional Preparatory Courses

- Preparatory course in matrices and calculus for students who need to rejuvenate their mathematical skills (August, 2023)
- Review course on basic probability and statistics concepts to solidify students' conceptual understanding (August, 2023)
- Workshop in R covering data handling, graphics, mathematical functions and some basic statistical techniques (August, 2023)
- Workshop in SAS for students who need to rejuvenate their skills in data management using SAS (August, 2023)
- Workshop on Python provides students a quick overview of the Python programming language. (August, 2023)

Students Testimonial

I am very thankful to have taken the MDASC programme and met so many supportive and knowledgeable professors. The program is comprehensive and covers the complete data science lifecycle. The courses did not only include theories, but also plenty of practices such as group projects and hands-on workshops. The programme equipped me with all the skillsets I needed to be a professional data scientist. I am now working in the Football Trading Department of HKJC and the lessons I learnt from the Master programme unquestionably facilitate my research and allow me to further enhance my models.

WONG Ho Yeung [MDASC Part-time Graduate 2022]
Senior Data Scientist (Sports Wagering), The Hong Kong Jockey Club

Participating in the MDASC program means a lot to me. It not only has courses in computer science, but also covers the subject of statistics. It made me realize the importance of data in my daily work and life. As a computer vision algorithm engineer, I deeply feel that data closed-loop has now become a key link in machine learning. The program equips me with basic skills in data mining and analysis, and helps me to better advance on the path of algorithm engineer. In addition, the MDASC program also provides a wealth of courses for students with different career paths to choose from, such as finance, data science, and cloud computing. I believe that every student who wants to understand what data science is and hope to develop in this area is suitable for joining this program.

CHEN Da [MDASC Full-time Graduate 2022]
Computer Vision Algorithm Engineer, Horizon Robotics

The MDASC programme is both academically challenging and commercially relevant allowing me to enhance my skills, experience and knowledge in various areas of data science. The learning experience through the MDASC is unparalleled. Not only do I learn from top professors in their fields, but also from talented and experienced classmates who come from different industries.

Specifically, the Data Science Project provides an excellent opportunity for the students to explore deeply in an interested field of data science and artificial intelligence or even commercialize their ideas. With the department's tremendous support and the programme's professional training, I have equipped with the knowledge, confidence and connections to start an artificial intelligence startup after graduation.

No matter where you come from or what you are looking for, I believe the MDASC programme will definitely open the door for a new career path for you in the future.

CHEUNG Ngai Yin [MDASC Part-time Graduate 2022]
Co-Founder, Mach Innovation

MDASC is a comprehensive master degree from theory in statistics to real world examples in computer science. With the elective subjects, I could not only study spatial data analysis for interest to know more about how to handle geographical data, but also could take the course for data visualization for work to illustrate the data analysis in an explainable and attractive way using Tableau to target audiences. With this program, I enhance my knowledge in statistical theory and learn more about the latest data science topics, such as cloud computing, deep learning, text mining and so on. It would be a good program for those who would like to know more about data science or work as a data scientist. After finishing the program, I am now more confident in explaining my models to my colleague and have a more clear direction on improving the accuracy, speed for current models and even deploy more complex machine learning models for my daily task.

CHOW Kwok Fung [MDASC Part-time Graduate 2021]
Senior Quantitative Developer, Nautilus Technology Limited

Coming from the finance field, I had very little experience in coding in my daily job. That being said, the MDASC program have provided me a great foundation to understand the theoretical aspect and real-life application around machine learning and data science. Having a structured curriculum combining with a final year capstone project, it was so helpful to glue the concepts together and apply it in the context which I believe is of great use for my future career. I would highly recommend this course to anyone who wants to take a well-organized program for Data Science.

LAM Chi Ho [MDASC Part-time Graduate 2021]
Associate Director, Barings Asset Management

The MDASC program has a well-balanced and flexible curriculum in Statistics and Computer Science. As a machine learning engineer, I work with large volume of data every day, which requires strong background in big data and machine learning techniques. The program equipped me with these core skills and built me a solid foundation for my career pursuit. Moreover, the flexibility of the curriculum and the diverse background of classmates help students to pursue any career path related to data, such as finance, data analyst, data scientist, etc. For me, the MDASC journey is a rewarding one, and I believe it will be a rewarding journey for anyone who believes the power of data.

WANG Kai [MDASC Full-time Graduate 2021]
Machine Learning Engineer, ByteDance

Examples of backgrounds of admitted students in recent years:

HKSAR Government departments/ units:

- Assistant Manager - Data Management
- Research Manager
- Statistical Assistant
- Data Analyst

Private companies:

- Director, Co-founder
- Technical Support Engineer
- Software Developer
- Software Engineer
- Data Engineer
- Actuarial Analyst
- Algorithm Engineer
- Business Analyst
- Senior Supervisor, Data Analyst
- Senior Associate, Project Management

Education profession:

- Research Assistant
- Tutor
- Subject Panel
- Data Analyst
- Assistant Researcher

Banking and finance profession:

- Associate, Marco Structuring
- Software Developer
- Senior Associate, Finance Investigations
- Data Scientist
- Fixed Income Associate
- Risk Management Engineer
- Senior Consultant
- Relationship Manager, Commercial Banking
- Financial Analytics Consultant (IT)
- Risk Actuary
- Assistant of Investment Counselor

Congregation

From 2021-22, two Congregations will be held annually (in July and December respectively) which is to align with international practices and to facilitate the convenience for students. Graduands who are eligible for graduation will be assigned to the Congregation nearest to the completion date of their studies for conferral of degree.

- Expected graduation time for normative study period
  Full Time *(1.5 year)* : Summer (July 2025)
  Part Time *(2.5 years)* : Summer (July 2026)

Tuition Fees

The full tuition fees for the programme is HK$286,200# for the 2023 intake. The fee shall normally be payable in three instalments over 1.5 years for full-time study or in five instalments over 2.5 years for part-time study. In addition, student is required to pay Caution Money (HK$350), which would be used to offset the graduation fee or, if the student is not graduating, be returned to the student.

 # Subject to approval

Target Students

It is a programme ideal for

- Those whose interest in high-level analytical skills straddles the disciplinary divide between statistics and computational analytics, and
- Those who wish to pursue further study in the field of data science after studying science, social sciences, engineering, medical sciences, information systems, computing and data analytics in their undergraduate studies.

Admission Requirements

- Applicants shall hold a Bachelor's degree with Honours or an equivalent qualification.
- Applicants shall have taken at least one university or post-secondary certificate course in each of the following three subjects (calculus and algebra, computer programming and introductory statistics) or related areas.
- Applicants shall fulfil the University Entrance Requirements.

Application Deadline

Main Round: 12 noon (GMT +8), December 14, 2022
Clearing Round: 12 noon (GMT +8), January 31, 2023

Programme Details

https://saasweb.hku.hk/programme/mdasc-index.php

Online Application

https://admissions.hku.hk/tpg/

Enquiries

Miss Jacey Yeung
Department of Statistics & Actuarial Science
Tel: (852) 3917 4152 Email: mdasc@hku.hk

Programme Director

Dr Marius HOFERT

MSc Syracuse; Dipl.-Math. oec., Dr. rer. nat. Ulm
Department of Statistics & Actuarial Science

STAFF LIST

Department of Statistics and Actuarial Science

Dr Y Cao	BS(Fudan); MS, PhD(Princeton)
Prof K C Cheung	BSc(ActuarSc), PhD HK; ASA
Dr L Feng	BS(Renmin U), PhD(Rutgers)
Dr K Han	PhD(HK)
Dr M Hofert	MSc Syracuse; Dipl.-Math. oec., Dr. rer. nat. Ulm
Dr C W Kwan	BSc, PhD HK
Dr E K F Lam	BA St. Thomas; MA New Brunswick; PhD HK
Dr A S M Lau	BEng(City), MSc(HK); PhD(CUHK)
Dr D Lee	BSc(ActuarSc), MPhil HK; PhD British Columbia; ASA
Prof S M S LEE	BA, PhD Cantab
Dr E A L Li	BSc HK; MEcon, PhD Syd
Prof G D Li	BSc, MSc Peking; PhD HK
Dr C Wang	PhD NUS
Dr K P Wat	BSc(ActuarSc), PhD HK; SFHEA; FSA; FASHK; CERA; FRM
Prof H L Yang	BSc Inner Mongolia; MMath Waterloo; PhD Alberta; ASA; HonFIA
Dr L Q Yu	BEng ZJU; PhD CUHK
Prof K C Yuen	BSc, MSc, PhD Calgary; ASA
Dr D Y Zhang	BSc Nankai; MSc, PhD NCSU
Dr M M Zhang	BS UCSB; MS, PhD UT Austin
Dr Z Q Zhang	BSc Nankai; MSc E China Normal; PhD HK
Dr K Zhu	BSc USTC; PhD HKUST

Department of Computer Science

Professor B Cautis	PhD INRIA & University of Paris South, Orsay
Dr B M Y Chan	MS UC San Diego, PhD HK
Dr V P S Chan	PhD HK
Professor F Y L Chin	PhD Princeton
Dr K P Chow	MA, PhD UC Santa Barbara
Dr T W Chim	PhD HK
Dr L Y K Choi	PhD HK
Professor W Y Chung	PhD University of Arizona
Dr T Luo	PhD HK
Professor A Montgomery	MBA, PhD University of Chicago
Dr D Schnieders	PhD HK
Professor M Sozio	PhD Sapienza University of Rome
Dr H F Ting	PhD Princeton
Dr R S W Yiu	PhD UC Berkeley
Professor S M Yiu	PhD HK

Faculty of Science, The University of Hong Kong
G/F Chong Yuet Ming Physics Building, Pokfulam Road, Hong Kong
Tel: (852) 3917 5287 Fax: (852) 2858 4620 Email: scitpg@hku.hk
https://www.scifac.hku.hk/pg/prospective/tpg

HKU Science