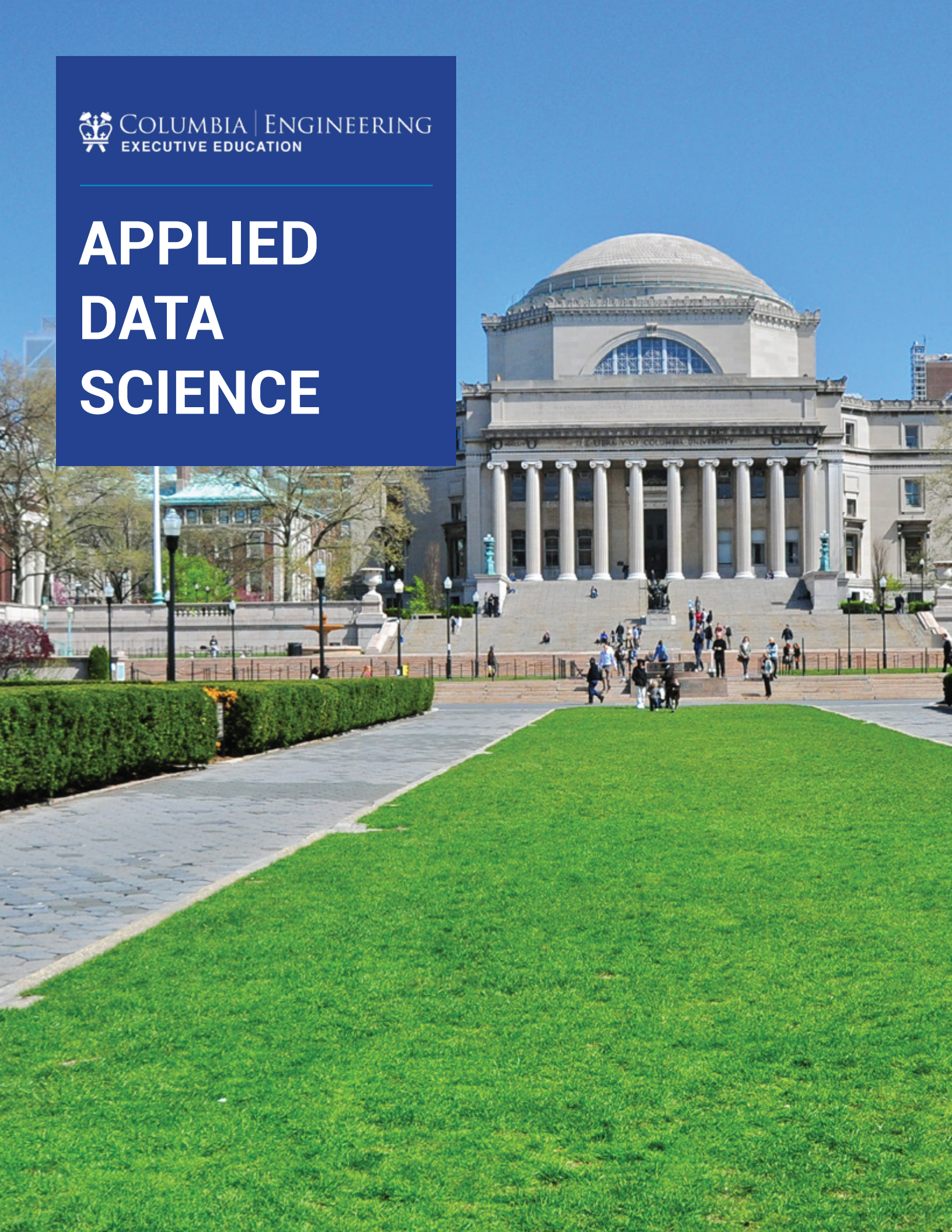
COLUMBIA | ENGINEERING
EXECUTIVE EDUCATION

APPLIED DATA SCIENCE



Overview

> “Information is the oil of the 21st century, and analytics is the combustion engine.”
>
> – Peter Sondergaard, Gartner Research

Data has been called the new global currency, and its meteoric rise is transforming entire industries—and driving the demand for practitioners who can wield its power. From health care and finance to entertainment, cybersecurity and beyond, the need for data scientists continues to grow in tandem with opportunities for career advancement within the field.

To help fill this talent gap and further the use of data science to solve real-world problems, Columbia Engineering Executive Education has partnered with EMERITUS to create the Applied Data Science course.

About Columbia Engineering

Columbia Engineering (Columbia University Fu Foundation School of Engineering and Applied Science), is committed to pushing the frontiers of knowledge and shaping discoveries to meet the needs of society. These aspirations have been fundamental since its early origins in 1864 as a school devoted to metallurgy and mining.

Over the years, Columbia's faculty and students have made remarkable contributions that have spurred technological and social progress. Today, Columbia carries the tradition of innovation as engineering transforms nearly every aspect of life.

Faculty at Columbia Engineering have won 10 Nobel Prizes in physics, chemistry, medicine, and economics.

Transcending Disciplines, Educating Leaders, Transforming Lives

10

Nobel Prizes in physics, chemistry, medicine, and economics

1864

The year of origin, we started pushing the frontiers of knowledge and shaping discoveries to meet the needs of society

Course Faculty

Vineet Goyal

Assistant Professor
Industrial Engineering and Operations Research

Professor Vineet Goyal has a Bachelor's degree in Computer Science from Indian Institute of Technology, Delhi and a Ph.D. from Carnegie Mellon University. Before coming to Columbia, he spent two years as a Postdoctoral Associate at the Operations Research Center at MIT. Professor Goyal is interested in the development of tractable approaches for dynamic optimization problems under uncertainty and their applications in electricity markets, revenue management and supply-chain and inventory management.

Costis Maglaras

David and Lyn Silfen Professor of Business

Costis Maglaras is the David and Lyn Silfen Professor of Business at Columbia University, and currently serves as the chair of the Decision, Risk & Operations division of the school. His research lies on the interface of stochastic modeling with operations management, with emphasis on stochastic networks, financial engineering, and quantitative pricing and revenue management. Costis received his BS in Electrical Engineering from Imperial College, London, and his MS and PhD in Electrical Engineering from Stanford University.

Hardeep Johar

Senior Lecturer of Industrial Engineering and Operations Research

Hardeep Johar received an M.A. in Economics from the Birla Institute of Technology and Science and is a Fellow of the Indian Institute of Management Calcutta. He received a Ph.D. in Information Systems from the Stern School of Business, New York University. Prior to joining Columbia, Johar has worked as a quantitative trader at Morgan Stanley, Credit Suisse and Deutsche Bank, at a tech startup (MSpoke), and has taught at NYU Stern School of Business and the Gabelli School of Business at Fordham University.

Course Leaders

Ian Jamieson

Course Leader, EMERITUS

Ian is a Research and Development Engineer in Boston and is involved in modeling neurophysiological processes and applying these models to engineering and other applications. He has developed and used a variety of artificial neural networks, Deep Learning and Machine Learning methods for applications in image processing (e.g. fMRI), medical diagnostics, and detecting biomedical anomalies, such as heart arrhythmia and neural conditions.

Ian's research interests also include Genetic Algorithms and Swarm Intelligence for producing novel solutions to classification methods. In addition to research, he has also taught Physics, Statistics, Mathematics, and astronomy/cosmology. Ian is a Member of the Institute of Physics in the UK and is a Chartered Physicist.

Kristen Kehrer

Course Leader, EMERITUS

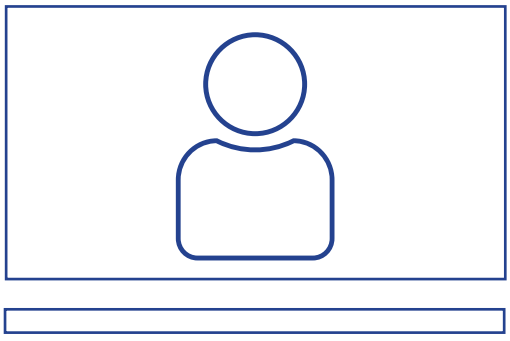
Kristen is #8 LinkedIn Global Top Voice 2018 – Data Science & Analytics. Since 2010, Kristen has been a data scientist across multiple industries, including the utilities, healthcare, and eCommerce. She finished a BS in Mathematics in 2004, and a Master's Degree in Applied Statistics. Prior to attaining her Master's Degree, she was a high school math teacher, and has always enjoyed tutoring, coaching, and mentoring.

**Course Leaders are subject to change*

Course Highlights

Our approach to this course is to teach the underlying concepts and statistics of Data Science.

Going beyond the theory, our approach invites participants into a conversation, where learning is facilitated by live subject matter experts and enriched by practitioners in the field of data science. We expect learners would be required to put in 6-8 hours per week.



✓ Live Online Teaching

✓ Faculty Video Lectures

✓ Quizzes

✓ Moderated Discussion Boards

✓ Q&A Sessions with Course Leaders

✓ Application Projects

Syllabus

Tools and Data Management

Module 1

Python Basics
Translating procedures into code

Module 2

Intermediate Python
Introduction to Data Structures

Module 3

Relational Databases
Where (most) data is stored

Module 4

SQL
Ubiquitous database formats/languages

Statistics & Exploratory Data Analysis

Module 5

Statistical Distributions
The shape of data

Module 6

Sampling
When you don't have all the data

Module 7

Hypothesis Testing
Answering questions about your data

Module 8

Data Analysis & Visualization
Using Python's NumPy for analysis

Module 9

Data Analysis & Visualization
Using Python's Pandas for data wrangling

Fundamentals of Machine Learning

Module 10

Text Mining
Automatic understanding of text

Module 11

Regression & Classification
Machine learning methods for prediction

Module 12

Clustering & Decision Trees
Machine learning methods for representation

Application Projects

Data Wrangling using CNC Mill Tool Wear Data

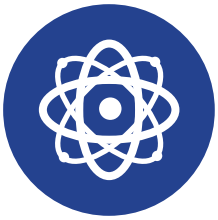
- Practice using Python's data framework to process and manipulate data with the CNC Mill Tool Wear dataset.
- Hone your data wrangling and munging skills using Python's pandas and NumPy libraries with the CNC Mill Tool Wear dataset.

Hypothesis Testing using Cancer Atlas Data

- Statistically test the impact of health factors in relation to cancer rates from around the globe.

Data Exploration using Lending Club Loan Data

- Use Python's NumPy library to explore and uncover insights in Lending Club's loan data.
- Using Python's powerful Pandas library to wrangle and munch Lending Club's loan data.

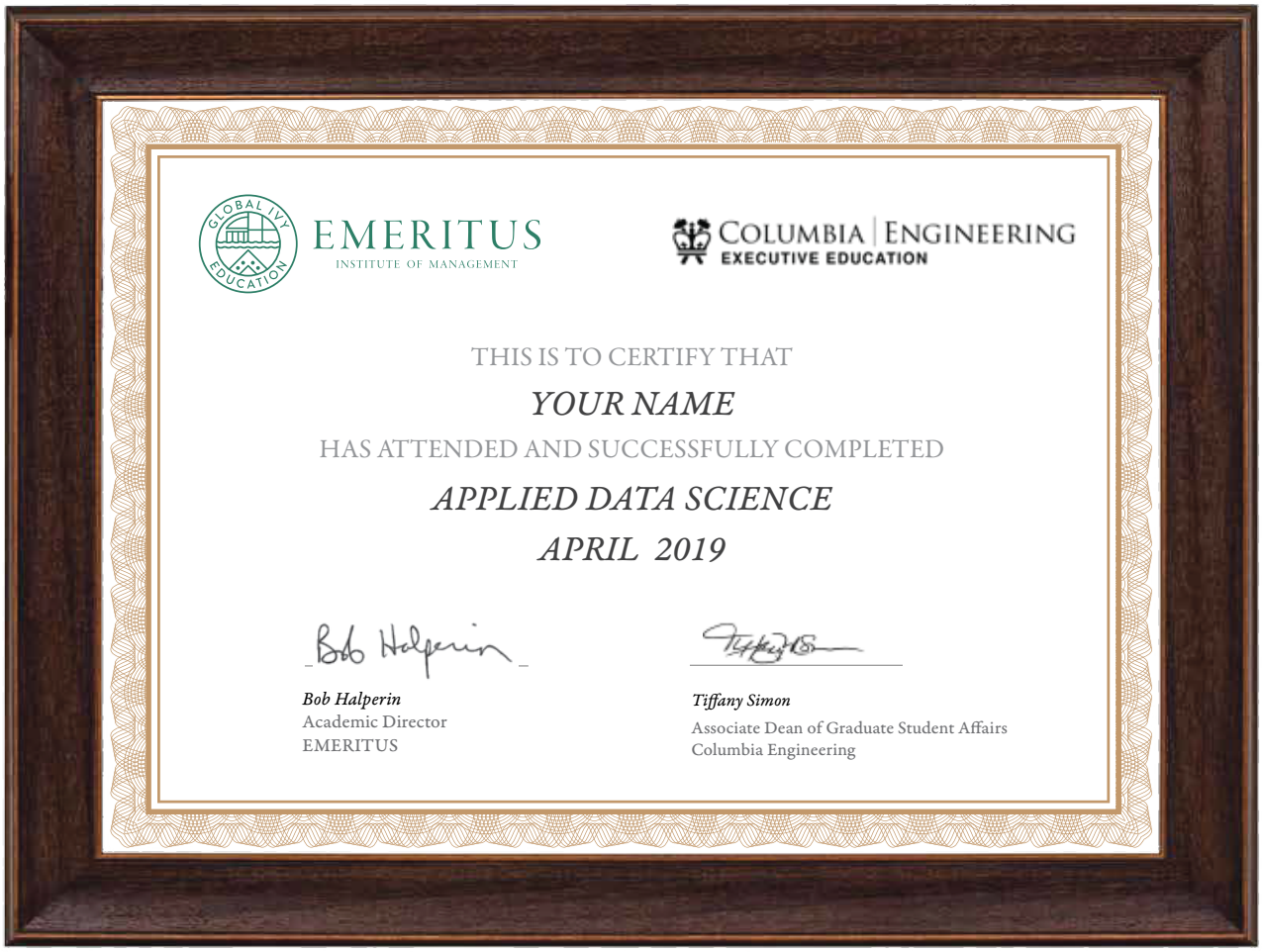


Natural Language Processing (NLP) implementation using Amazon product reviews

- Implement Natural Language Processing (NLP) techniques to automate the understanding of product reviews from Amazon.

Certificate

Upon successful completion of the course, participants will receive a verified digital certificate from EMERITUS in collaboration with Columbia Engineering Executive Education.



Course Fee and Duration

USD 1,400
3 Months, Online

Singapore residents who wish to enroll for this course will be charged GST.

Pre-Requisites

The course requires an undergraduate knowledge of statistics, (descriptive statistics, regression, sampling distributions, hypothesis testing, interval estimation etc.) linear algebra and probability.

Familiarity with Python or any programming language is required.

Assignments/application projects which require programming will be done using the Python programming language.

About EMERITUS

Columbia Engineering Executive Education is collaborating with online education provider EMERITUS Institute of Management to offer a portfolio of high-impact online courses. These courses leverage Columbia's thought leadership in management practice developed over years of research, teaching, and practice.

By collaborating with EMERITUS, we are able to broaden access beyond our on-campus offerings in a collaborative and engaging format that stays true to the quality of Columbia.

EMERITUS' approach to learning is formulated on a cohort-based design to maximize peer-to-peer sharing and includes live teaching with world-class faculty and hands-on project-based learning. In the last year, more than 10,000 students from over 120 countries have benefited professionally from EMERITUS.

- ✓ **Online collaboration with EMERITUS**
- ✓ **Collaborative and engaging format**
- ✓ **Live teaching, hands-on project-based learning**
- ✓ **More than 10,000 students from over 120 countries have benefited**

www.emeritus.org

E-mail: info@emeritus.org

We attempt to respond to queries in 24 hours or less. However, over weekends and holidays, our responses may take up to 72 hours.

EMERITUS Institute of Management | Council for Private Registration No. 201510637C | Period: 29 March 2018 — 28 March 2022